If the triangle has an obtuse angle, say at A, then the corresponding term NP has the negative sign in this relation:

$$MN + MP - NP = 2\Delta/R.$$

The theorems which follow are subject to similar modification.

. . . .

§23. Since $AN = c \cos \alpha$, $AP = b \cos \alpha$, and $\frac{1}{2}bc \sin \alpha = \Delta$, it follows that $\triangle ANP = \Delta.\cos^2 \alpha$; likewise $\triangle BMP = \Delta.\cos^2 \beta$ and $\triangle CMN = \Delta.\cos^2 \lambda$; hence

$$\triangle MNP = \Delta(1 - \cos^2 \alpha - \cos^2 \beta - \cos^2 \gamma);$$

but since $\cos^2 \gamma = (\cos \alpha \cos \beta - \sin \alpha \sin \beta)^2$ and $\sin^2 \alpha \sin^2 \beta = 1 - \cos^2 \alpha - \cos^2 \beta + \cos^2 \alpha \cos^2 \beta$, then

$$\cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma = 1 - 2 \cos \alpha \cos \beta \cos \gamma,$$

and accordingly

$$\triangle MNP = 2\Delta \cos \alpha \cos \beta \cos \gamma = \frac{(-a^2 + b^2 + c^2)(a^2 - b^2 + c^2)(a^2 + b^2 - c^2)}{4a^2b^2c^2}\Delta.$$

§24. If ρ designates the radius of the inscribed circle of triangle MNP and $\rho^{(1)}$, $\rho^{(2)}$, $\rho^{(3)}$ those of the escribed circles, then, by virtue of §2 and §23, we have for the acute triangle ABC:

$$\rho = \frac{4\Delta \cos \alpha \cos \beta \cos \gamma}{a \cos \alpha + b \cos \beta + c \cos \gamma}$$

$$\rho^{(1)} = \frac{4\Delta \cos \alpha \cos \beta \cos \gamma}{-a \cos \alpha + b \cos \beta + c \cos \gamma}$$

[and if angle α is obtuse, these equations are modified by changing the sign of each term containing $\cos \alpha$.]

Now since in general, by §19, $a \cos \alpha + b \cos \beta + c \cos \gamma = 2\Delta/R$, we have for the acute triangle ABC:

$$\rho = 2R \cos \alpha \cos \beta \cos \gamma$$

and on the other hand, if A is obtuse,

$$\rho^{(1)} = -2R \cos \alpha \cos \beta \cos \gamma.$$

. . . .

26. The radius of the circle circumscribed about the triangle MNP is equal to

$$\frac{MN.MP.NP}{4\, \Delta\, MNP} = \frac{abc \cos \alpha \cos \beta \cos \gamma}{8\Delta \cos \alpha \cos \beta \cos \gamma} = \tfrac{1}{2}R$$

that is, to half the radius of the circle circumscribed about the triangle.

. . . .

§32. Since angle AOP equals ABC, then $AO = \frac{AP}{\sin\beta}$, and since $AP = b\cos\alpha$, and $\sin\beta = \frac{b}{2R}$, therefore:

$AO = 2R\cos\alpha$; similarly $BO = 2R\cos\beta$ and $CO = 2R\cos\gamma$; hence

$$AO + BO + CO = 2R(\cos\alpha + \cos\beta + \cos\gamma)$$

[and by substitution of the formulas for the cosines and algebraic reduction, we find for any acute triangle]

$$\cos\alpha + \cos\beta + \cos\gamma = \frac{r+R}{R},\ AO + BO + CO = 2(r+R)$$

[If the triangle has an obtuse angle, *e. g.*, at C, then]

$$AO + BO - CO = 2(r+R).$$

. . . .

§35. We have $OM = BO\cos\gamma$, and since by §32 $BO = 2R\cos\beta$, therefore

$$OM = 2R\cos\beta\cos\gamma;$$

similarly $ON = 2R\cos\alpha\cos\gamma$, and $OP = 2R\cos\alpha\cos\beta$. If one multiplies these expressions respectively by those of AO, BO, CO (§32), then, since $\cos\alpha\cos\beta\cos\gamma = \frac{\rho}{2R}$ (§24)

$$AO.OM = BO.ON = CO.OP = 2\rho R$$

That is, *the point of intersection of the three perpendiculars of triangle ABC divides each into two parts, whose rectangle equals double the rectangle of the radius of the circle inscribed in triangle MNP and that of the circle circumscribed about triangle ABC.*

. . . .

CHAPTER III

On the Center of the Circle, Which Is Circumscribed about a Triangle

§45. If K is the center of the circle circumscribed about a triangle ABC, and if perpendiculars Ka, Kb, Kc are dropped from this point on the sides, BC, CA, AB; then if we draw AK, $Kc = AK\cos AKc$; and because $AK = R$ and angle AKc equals ACB, therefore

$$Kc = R\cos\gamma;$$

and similarly $Kb = R\cos\beta$ and $Ka = R\cos\alpha$. If we compare these expressions with those found in §32 for AO, BO, CO,... we have at once

$$AO = 2Ka,\ \ BO = 2Kb,\ \ CO = 2Kc,$$

In any triangle the distance from the center of the circumscribed circle to any side is half the distance from the common point of the altitudes to the opposite vertex.

. . . .

CHAPTER IV

Determination of the Relative Positions of the Points Previously Discussed. [Bestimmung der Gegenseitigen Lage der Vornehmsten bisher betrachteten Punkte.]

§49. If K and S are the centers of the circles circumscribed and inscribed to the triangle ABC, and perpendiculars Kc and SF are dropped from these to the side AB, then

$$\overline{KS}^2 = (Ac - AF)^2 + (SF - Kc)^2$$

Now we have $Ac = \frac{1}{2}c$ and $AF = \frac{1}{2}(-a + b + c)$, whence:

$$Ac - AF = \tfrac{1}{2}(a - b);$$

further, because (§2)

$$SF = \frac{2\Delta}{a + b + c},$$

and (§45)

$$Kc = \frac{c(a^2 + b^2 - c^2)}{8\Delta},$$

therefore

$$SF - Kc = \frac{(-a + b + c)(a - b + c)(a + b - c) - c(a^2 + b^2 - c^2)}{8\Delta}$$

If now we substitute in the above expression for $\overline{KS}^2$, after the necessary reductions we have

$$\overline{KS}^2 = \frac{a^2b^2c^2 - abc(-a + b + c)(a - b + c)(a + b - c)}{16\Delta^2},$$

whence by means of the known values of the radii r and R we arrive at the result:

$$\overline{KS}^2 = R^2 - 2rR$$

In any triangle the square of the distance between the centers of the in- and circumscribed circles equals the square of the radius of the circumscribed circle, diminished by twice the rectangle of this radius and the radius of the inscribed circle.

[By similar methods we find that if S' is the center and r' the radius of an escribed circle,

$$\overline{KS'}^2 = R^2 + 2r'R$$[1]

. . . .

§51, 53. [By exactly the same methods, we derive the relations]

$$\overline{OS}^2 = 2r^2 - 2\rho R$$
$$\overline{KO}^2 = R^2 - 4\rho R$$

§54. If L is the center of the circle circumscribed about the circle MNP, whose radius (§26) has been found to be $\frac{1}{2}R$, and if OL is drawn, then since O is also the center of the inscribed circle of triangle MNP, then (§49) $\overline{OL}^2 = \frac{1}{4}R^2 - \rho R$, and since we have just seen that $\overline{KO}^2 = R^2 - 4\rho R$, it follows that $\overline{KO}^2 = 4\overline{OL}^2$, or

$$KO = 2OL.$$

[If the triangle is obtuse, a slight modification of the proof leads to the same result, viz:]

In any triangle the common point of the altitudes is twice as far from the center of the circumscribed circle as from the center of the circle through the feet of the altitudes.

§55. If perpendiculars LJ, LH are dropped from the center L on the lines AB, CP, $LJ = PH$, and since H is a right angle, it is known that in triangle OPL, $PH = \dfrac{-\overline{OL}^2 + \overline{OP}^2 + \overline{LP}^2}{2OP}$. But $LP = \frac{1}{2}R$, and (§54) $\overline{OL}^2 = \frac{1}{4}R^2 - \rho R$; further (§35, 45) $OP.Kc = \rho R$, whence $\overline{LP}^2 - \overline{OL}^2 = OP.Kc$. If this expression is substituted in $PH = LJ$, the result is

$$LJ = \tfrac{1}{2}(OP + Kc).$$

From this property it follows at once that the points O, L, K are on one and the same line, and the theorem shines out (erhellet):

In any triangle the center of the circumscribed circle, the common point of the altitudes, and the center of the circle through feet of the altitudes lie on one and the same straight line, whose mid-point is the last named point.

§56. Because the point L, then, lies at the center of the line KO, therefore the point J is also the center of the line Pc, whence

[1] In a historical note, the theorem is attributed to Euler. The history of this theorem has been investigated in detail by Mackay, *Proceedings of the Edinburgh Math. Society*, V. 1886–7, p. 62.

it follows that $Lc = LP = \frac{1}{2}R$; and similarly on each of the other sides AC, BC. Thus we have the theorem:

In any triangle the circle which passes through the feet of the altitudes also cuts the sides at their mid-points.

§57. If the line LS is drawn, we know that in triangle KOS, since L is the mid-point of KO, $2\overline{SL}^2 + 2\overline{OL}^2 = \overline{KS}^2 + \overline{OS}^2$. If we substitute in this equation the values of the squares of OL KS, OS, as found in 54, 49, 51, thus there comes

$$\overline{LS}^2 = \tfrac{1}{4}R^2 - rR + r^2 = (\tfrac{1}{2}R - r)^2,$$

or:

$$LS = \tfrac{1}{2}R - r.$$

Similarly, setting a, b, c in turn negative,

$$LS' = \tfrac{1}{2}R + r',\ \ LS'' = \tfrac{1}{2}R + r'',\ \ LS''' = \tfrac{1}{2}R + r'''.$$

Since now (§26)$\frac{1}{2}R$ is the radius of the circle circumscribed about triangle MNP, we deduce from a known property of circles which are tangent, the following theorem:

The circle which passes through the feet of the altitudes of a triangle touches all four of the circles which are tangent to the three sides of the triangle and specifically, it touches the inscribed circle internally and the escribed circles externally.

WILLIAM JONES

The First Use of π for the Circle Ratio

(Selections Made by David Eugene Smith from the Original Work.)

William Jones (1675–1749) was largely a self-made mathematician. He had considerable genius and wrote on navigation and general mathematics. He edited some of Newton's tracts. The two passages given below are taken from the *Synopsis Palmariorum Matheseos: or, a New Introduction to the Mathematics*, London, 1706. The work was intended "for the Vse of some Friends, who had neither Leisure, Conveniency, nor, perhaps, Patience, to search into so many different Authors, and turn over so many tedious volumes, as is unavoidably required to make but a tolerable Progress in the Mathematics." It was a very ingenious compendium of mathematics as then known. The symbol π first appears on page 243, and again on p. 263. The transcendence of π was proved by Lindemann in 1882. For the transcendence of e, which proved earlier (1873), see page 99.

Taking a as an arc of 30°, aud t as a tangent in a figure given, he states (p. 243):

$$6a, \text{ or } 6 \times t - \frac{1}{3}t^2 + \frac{1}{5}t^5, \text{ \&c.} = \frac{1}{2} \text{ Periphery } (\pi) \ldots$$

Let

$$\alpha = 2\sqrt{3},\ \beta = \frac{1}{3}\alpha,\ \gamma = \frac{1}{3}\beta,\ \delta = \frac{1}{3}\gamma, \text{ \&c.}$$

Then

$$\alpha - \frac{1}{3}\beta + \frac{1}{5}\gamma - \frac{1}{7}\delta + \frac{1}{9}\epsilon, \text{ \&c.} = \frac{1}{2}\pi,$$

or

$$\alpha - \frac{1}{3}\frac{3\alpha}{9} + \frac{1}{5}\frac{\alpha}{9} - \frac{1}{7}\frac{3\alpha}{9^2} + \frac{1}{9}\frac{\alpha}{9^2} - \frac{1}{11}\frac{3\alpha}{9^3} + \frac{1}{13}\frac{\alpha}{9^3}, \text{ \&c.}$$

Theref. the (Radius is to ½ Periphery, or) Diameter is to the Periphery, as 1,000, &c to 3.141592653 . 5897932384 . 6264338327 . 9502884197 . 1693993751 . 0582097494 . 4592307816 . 4062862089 . 9862803482 . 5342117067. 9+ True to above a 100 Places; as Computed by the accurate and Ready Pen of the Truly Ingenious Mr. *John Machin*.

On p. 263 he states:

There are various other ways of finding the *Lengths*, or *Areas* of particular *Curve Lines*, or *Planes*, which may very much facili-

tate the Practice; as for Instance, in the *Circle*, the *Diameter* is to *Circumference* as 1 to

$$\frac{16}{3} - \frac{4}{239} - \frac{1}{3}\frac{16}{5^3} - \frac{4}{239^3} + \frac{1}{5}\frac{16}{5^5} - \frac{4}{239^5} -, \&c. = 3.14159, \&c. = \pi \ldots$$

Whence in the *Circle*, any one of these three, α, c, d, being given, the other two are found, as, $d = c \div \pi = \overline{\alpha \div \frac{1}{4}\pi}|^{\frac{1}{2}}$, $c = d \times \pi = \overline{\alpha \times 4\pi}|^{\frac{1}{2}}$, $\alpha = \frac{1}{4}\pi d^2 = c^2 \div 4\pi$.

GAUSS

On the Division of a Circle into n Equal Parts

(Translated from the Latin by Professor J. S. Turner, University of Iowa, Ames, Iowa.)

Carl Friedrich Gauss (1777–1855) was a student at Göttingen from 1795 to 1798, and during this period he conceived the idea of least squares, began his great work on the theory of numbers (*Disqvisitiones arithmeticae*, Leipzig, 1801), and embodied in the latter his celebrated proposition that a circle can be divided into n equal parts for various values of n not theretofore known. This proposition is considered in the *Disqvisitiones*, pages 662–665. From this edition the following translation has been made, the portion selected appearing in sections 365 and 366.

For further notes upon Gauss and his works see pages 107 and 292.

(365.) We have therefore, by the preceding investigations, reduced the division of the circle into n parts, if n is a prime number, to the solution of as many equations as there are factors into which $n - 1$ can be resolved, the degrees of these equations being determined by the magnitude of the factors. As often therefore as $n - 1$ is a power of the number 2, which happens for these values of n: 3, 5, 17, 257, 65537 *etc.*, the division of the circle is reduced to quadratic equations alone, and the trigonometric functions of the angles $\frac{P}{n}, \frac{2P}{n}$ *etc.* can be expressed by square roots more or less complicated (according to the magnitude of n); hence in these cases the division of the circle into n parts, or the description of a regular polygon of n sides, can evidently be effected by geometrical constructions. Thus for example for $n = 17$, by arts. 354, 361, the following[1] expression is easily derived for the cosine of the angle $\frac{1}{17}P$:

$$-\frac{1}{16} + \frac{1}{16}\sqrt{17} + \frac{1}{16}\sqrt{(34 - 2\sqrt{17})} - \frac{1}{8}\sqrt{(17 + 3\sqrt{17} - \sqrt{(34 - 2\sqrt{17})} - 2\sqrt{(34 + 2\sqrt{17})})};$$

[1] [An elegant presentation of Gauss's method will be found on p. 220 of Casey's *Plane Trigonometry* (Dublin, 1888), where, however, the last terms of equations (550), (551), (552) should be c_1, c_2, b_2 respectively.]

the cosines of the multiples of this angle have a similar form, but the sines have one more radical sign. Truly it is greatly to be wondered at, that, although the geometric divisibility of the circle into three and five parts was already known in the times of Euclid, nothing has been added to these discoveries in the interval of 2000 years, and all geometers have pronounced it as certain, that beyond the divisions referred to and those which readily follow, namely divisions into 15, 3.2^μ, 5.2^μ, 15.2^μ and also into 2^μ parts, no others can be effected by geometrical constructions. Moreover it is easily proved, if the prime number n is equal to $2^m + 1$, that the exponent m cannot involve other prime factors than 2, and so must be either equal to 1 or 2 or a higher power of 2; for if m were divisible by any odd number ζ (greater than unity), and $m = \zeta\eta$, $2^m + 1$ would be divisible by $2^\eta + 1$, and therefore necessarily composite. Consequently all values of n by which we are led to none but quadratic equations are contained in the form $2^{2^\nu} + 1$; thus the 5 numbers 3, 5, 17, 257, 65,537 result by setting $\nu = 0$, 1, 2, 3, 4 or $m = 1$, 2, 4, 8, 16. By no means for all numbers contained in that form, however, can the division of the circle be performed geometrically, but only for those which are prime numbers. Fermat indeed, misled by induction, had affirmed that all numbers contained in that form are necessarily primes; but the celebrated Euler first remarked that this rule is erroneous even for $\nu = 5$, or $m = 32$, the number $2^{32} + 1 = 4294967297$ involving the factor 641.

But as often as $n - 1$ involves other prime factors than 2, we are led to higher equations; namely to one or more cubics when 3 is found once or more frequently among the factors of $n - 1$; to equations of the 5^{th} degree when $n - 1$ is divisible by 5 *etc.*, AND WE CAN DEMONSTRATE WITH ALL RIGOR THAT THESE HIGHER EQUATIONS CAN IN NO WAY BE EITHER AVOIDED OR REDUCED TO LOWER, although the limits of this work do not permit this demonstration to be given, which nevertheless we effected since a warning must be given lest anyone may still hope to reduce to geometrical constructions other divisions beyond those which our theory furnishes, for example divisions into 7, 11, 13, 19 *etc.* parts, and waste his time uselessly.

(366.) If the circle is to be divided into a^α parts, where a denotes a prime number, this can clearly be effected geometrically when $a = 2$, but for no other value of a, provided $a > 1$; for then besides those equations which are required for the division into a

parts it is also necessary to solve $\alpha - 1$ others of the a^{th} degree; moreover these can in no way be either avoided or depressed. Consequently the degrees of the necessary equations can be ascertained generally (evidently also for the case where $\alpha = 1$) from the prime factors of the number $(a - 1)a^{\alpha-1}$.

Finally, if the circle is to be divided into $N = a^\alpha b^\beta c^\lambda \ldots$ parts, a, b, c *etc.* denoting unequal prime numbers, it suffices to effect the divisions into a^α, b^β, c^γ *etc.* parts (art. 336); and therefore, to ascertain the degrees of the equations required for this purpose, it is necessary to examine the prime factors of the numbers $(a - 1)a^{\alpha-1}$, $(b - 1)b^{\gamma-1}$, $(c - 1)c^{\gamma-1}$ *etc.*, or which amounts to the same thing, of the product of these numbers. It may be observed that this product expresses the number of numbers prime to N and less than N (art. 38). Therefore the division is effected geometrically only when this number is a power of two; indeed when it involves prime factors other than 2, for instance p, p' *etc.*, equations of degree p, p' *etc.* can in no way be avoided. Hence it is deduced generally that, in order that a circle may be geometrically divisible into N parts, N must be *either* 2 or a higher power of 2, *or* a prime number of the form $2^m + 1$, *or* the product of several such prime numbers, *or* the product of one or more such prime numbers into 2 or a higher power of 2; or briefly, it is necessary that N should involve neither any odd prime factor which is not of the form $2^m + 1$, nor even any prime factor of the form $2^m + 1$ more than once. The following 38 such values of N are found below 300: 2, 3, 4, 5, 6, 8, 10, 12, 15, 16, 17, 20, 24, 30, 32, 34, 40, 48, 51, 60, 64, 68, 80, 85, 96, 102, 120, 128, 136, 160, 170, 192, 204, 240, 255, 256, 257, 272.

SACCHERI

On Non-Euclidean Geometry[1]

(Translated from the Latin by Professor Henry P. Manning, Brown University, Providence, R. I.)

Geronimo Saccheri was born in 1667 and died in 1733. He was a Jesuit and taught in two or three of the Jesuit colleges in Italy. His chief work, published about the time of his death, is an attempt to prove Euclid's parallel postulate as a theorem by showing that the supposition that it is not true leads to a contradiction. The path to his "contradiction" consists of a series of propositions which actually constitute the main part of the elementary non-Euclidean geometry, published in this way about a hundred years before it was published as such.

The final discovery of the non-Euclidean geometry was not based on the work of Saccheri. Neither Lobachevsky nor Bolyai seems to have ever heard of him. But Saccheri is the most important figure in the preparation for this discovery in the period that precedes it, and after his relation to it was pointed out in 1889 his work took its place as standing at the head of the literature of the subject.

Euclid Freed from Every Flaw[2]

Book I

in which is demonstrated: Any two straight lines lying in the same plane, on which a straight line makes the two interior angles on the same side less than two right angles, will at length meet each other on the same side if they are produced to infinity.

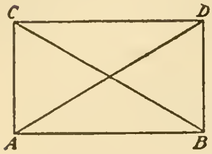
---

[1] On the general subject of non-Euclidean geometry, including biographical notes, see Engel and Stäckel, *Die Theorie der Parallellinien* (Leipzig, 1895); *Urkunden zur Geschichte der Nichteuklidischen Geometrie* (Leipzig, 1898, 1913).

[[2] This book was written in Latin and the Latin text has been published along with the translation by Professor Halsted (Chicago, 1920). We have checked this with the German translation given in *Die Theorie der Parallellinien* by Engel and Stäckel, Leipzig, 1895, pages 41–135. The text itself of which we are translating a part is preceded by a "Preface to the Reader" and a summary of the contents "added in place of an index."]

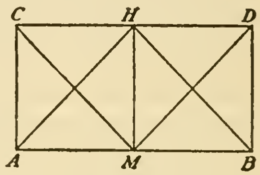
Part I

Proposition I.—If two equal straight lines, AC and BD, make equal angles on the same side with a line AB, I say that the angles with the joining line CD will also be equal.



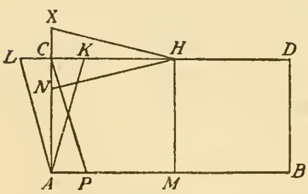
Proof.[1]—Let A and D be joined, and C and B. Then let the triangles CAB and DBA be considered. It follows (I, 4)[2] that the bases CB and AD will be equal. Then consider the triangles ACD and BDC. It follows (I, 8) that the angles ACD and BDC will be equal. Q. E. D.[3]

Proposition II.—In the same quadrilateral ABCD[4] *let the sides AB and CD be bisected at the points M and H. I say that the angles with the joining line MH will then be right angles.*



Proof.—Let the joining lines AH and BH be drawn, also CM and DM. Since in this quadrilateral the angles A and B are given equal, and also (from the preceding) C and D, it follows from I, 4 (since also the equality of the sides is known) that in the triangles CAM and DBM the bases CM and DM will be equal; also in the triangles ACH and BDH the bases AH and BH. Therefore, from a comparison of the triangles CHM and DHM, and again of the triangles AMH and BMH, it will follow (I, 8) that the angles in these at the points M and H will be equal to each other, and so right angles. Q. E. D.

Proposition III.—If two equal straight lines AC and BD stand perpendicularly to a straight line AB, I say that the joining line CD will be equal to, or less than, or greater than AB, according as the angles with the same CD are right or obtuse or acute.



Proof of the first part. Each angle C and D being a right angle, if possible let one of those, say CD, be greater than the other, AB. On DC

[1] ["Demonstratur," *It is proved.*]

[2] [..."ex quarta primi." This is a reference to Euclid. For such references we shall give only the numbers of the book and proposition.]

[3] ["Quod erat demonstrandum," written out in full in the original as published with Halsted's translation.]

[4] [Literally, "The uniform quadrilateral remaining."]

let a portion DK be taken equal to BA, and let A and K be joined. Since therefore on BD stand two equal perpendicular lines BA and DK the angles BAK and DKA will be equal (1).[1] But this is absurd, since the angle BAK is by construction less than the assumed right angle BAC, and the angle DKA is an exterior angle by construction, and therefore (I, 16) greater than the interior opposite angle DCA, which is supposed to be a right angle. Therefore neither of the given lines DC and BA is greater than the other if the angles with the joining line CD are right angles, and therefore they are equal to each other. Q. E. D. for the first part.

Proof of the second part. But if the angles with the joining line CD are obtuse, let AB and CD be bisected at the points M and H, and let M and H be joined. Since therefore on the straight line MH stand two perpendicular lines AM and CH (from what precedes) and we have with the joining line AC the right angle at A, CH will not be equal to AM (1) since the angle at C is not a right angle. But neither will it be greater: otherwise, taking on HC a portion KH equal to AM, we shall have equal angles with the joining line AK (1). But this is absurd as above. For the angle MAK is less than a right angle and the angle HKA is greater than the obtuse angle HCA which is interior and opposite (I, 16). It results therefore that CH, while the angles with the joining line CD are obtuse, is less than AM, and therefore the double of the former, CD, is less than the double of the latter, AB. Q. E. D. for the second part.

Proof of the third part. But, finally, if the angles with the joining line CD were acute, the perpendicular MH being drawn as before, we proceed thus: Since on the line MH stand perpendicularly two straight lines AM and CH, and with the joining line AC there is a right angle at A, the line CH will not be equal to AM, since there is lacking a right angle at C. But neither will it be less: otherwise, if on HC produced we take $HL = AM$, the angles formed with the joining line AL will be equal (as above). But this is absurd. For the angle MAL is by construction greater than the angle MAC supposed a right angle, and the angle HLA is by construction interior and opposite, and so less than the exterior angle HCA (I, 16), which is supposed acute. It remains, therefore, that CH, while the angles with the joining line CD are acute, is greater

[1] [... "ex prima hujus." These are references to previous theorems, which we will indicate simply by putting the number in the parentheses.]

than AM, and so CD, the double of the former, is greater than AB, the double of the latter. Q. E. D. for the third part.

Thus it follows that the joining line CD will be equal to, or less than, or greater than AB, according as the angles with the same CD are right or obtuse or acute. Q. E. D.

Corollary 1.—*Hence in every quadrilateral containing three right angles and one obtuse or acute, the sides adjacent to the angle which is not a right angle are less, respectively, than the opposite sides if the angle is obtuse, but greater if it is acute.* For it has already been demonstrated of the side CH with respect to the opposite side AM, and in a similar way it is shown of the side AC with respect to the opposite side MH. For as the lines AC and MH are perpendicular to AM, they cannot be equal to each other (1), because of the unequal angles with the joining line CH. But neither (in the hypothesis of the obtuse angle at C) can a certain portion AN of AC be equal to MH, than which certainly AC is greater; otherwise (1) the angles with the joining line HN would be equal which is absurd as above. But again (in the hypothesis of the acute angle at C) if we wish that a certain AX, taken on AC produced, shall be equal to MH, than which certainly AC is smaller, the angles with the joining line HX will be equal for the same reason, which is absurd in the same way as above. It remains therefore that in the hypothesis indeed of the obtuse angle at the point C the side AC will be less than the opposite side MH, but in the hypothesis of the acute angle it will be greater. Q. E. I.[1]

Corollary 2.—*By much more will CH be greater than any portion of AM,* as say PM, since the joining line CP makes a more acute angle with CH on the side towards the point H, and an obtuse angle (I, 16) with PM towards the point M.

Corollary 3.—*Again it follows that all these statements are true if the perpendiculars AC and BD are of a certain finite length fixed by us, or are supposed to be infinitely small.* This indeed ought to be noted in the rest of the propositions that follow.

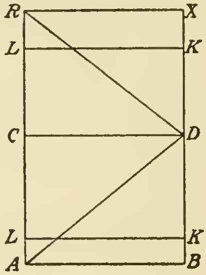
Proposition IV.—*But conversely (in the figure of the preceding proposition), the angles with the joining line CD* will be *right or obtuse or acute according as CD is equal to, or less than, or greater than, the opposite AB.*

Proof.—For if the line CD is equal to the opposite AB, and nevertheless the angles with the same are obtuse or acute, already such angles will prove it (from the preceding) not equal to, but less

[1] ["Quod erat intentum," *What was asserted.*]

than, or greater than the opposite AB, which is absurd, contrary to hypothesis. The same holds in the other cases. It stands therefore that the angles with the joining line CD are right or obtuse or acute according as the line CD is equal to, or less than, or greater than, the opposite AB. Q. E. D.

Definitions.—Since (from 1) a straight line joining the extremities of equal lines standing perpendicularly to the same line (which we shall call base) makes equal angles with them, therefore there are three hypotheses to be distinguished in regard to the nature of these angles. And the first indeed I will call *the hypothesis of the right angle,* but the second and third I will call *the hypothesis of the obtuse angle* and *the hypothesis of the acute angle.*



Proposition V.—The hypothesis of the right angle, if true in a single case, is always in every case the only true hypothesis.

Proof.—Let the joining line CD make right angles with any two equal lines, AC and BD, standing perpendicularly to any AB. CD will be equal to AB (3). Take on AC and BD produced the two CR and DX, equal to AC and BD, and join R and X. We shall easily show that the joining line RX will be equal to AB and the angles with it right angles; and first indeed by superposition of the quadrilateral $ABDC$ upon the quadrilateral $CDXR$, with the common base CD. But then we can proceed more elegantly thus: join A and D and R and D. It follows (I, 4) that the triangles ACD and RCD will be equal, the bases AD and RD, and also the angles CDA and CDR and, because they are equal remainders to a right angle, ADB and RDX. Wherefore again (from the same I, 4) will be equal in the triangles ADB and RDX the base AB to the base RX. Therefore (from the preceding) the angles with the joining line RX will be right angles, and therefore we shall persist in the same hypothesis of the right angle.

And since the length of the perpendiculars can be increased to infinity on the same base, with the hypothesis of the right angle always holding, it must be proved that the same hypothesis will remain in the case of any diminution of the same perpendiculars, which is proved as follows.

Take in AR and BX any two equal perpendiculars AL and BK, and join L and K. Even if the angles with the joining line are not right angles, yet they will be equal to each other (1). They

will therefore be obtuse on one side, say towards AB, and towards RX acute, as the angles at each of these points are equal to two right angles (I, 13). But it follows also that LR and KX are perpendiculars equal to each other standing on RX. Therefore (3) LK will be greater than the opposite RX and less than the opposite AB.

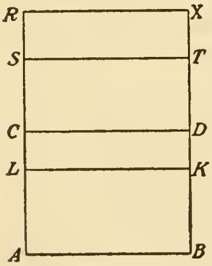
But this is absurd, since AB and RX have been shown equal. Therefore the hypothesis of the right angle will not be changed under any diminution of the perpendiculars while the given AB remains the base.

But neither will the hypothesis of the right angle be changed under any diminution or greater amplitude of the base, since it is evident that any perpendicular BK or BX can be considered as base, and so in turn AB and the equal opposite line KL or XR can be considered as the perpendicular.

It follows therefore that the hypothesis of the right angle, if true in any case, is always in every case the only true hypothesis. Q. E. D.

Proposition VI.—The hypothesis of the obtuse angle, if it is true in one case, is always in every case the only true hypothesis.

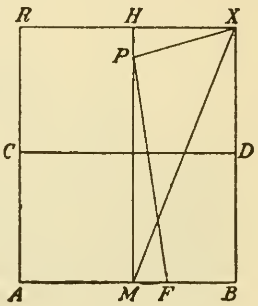
Proof.—Let the joining line CD make obtuse angles with any two equal perpendiculars AC and BD standing on any straight line AB. CD will be less than AB (3). Take on AC and BD produced any two portions CR and DX, equal to each other, and join R and X. Now I seek in regard to the angles with the joining line RX, which will be equal to each other (1). If they are obtuse we have the theorem as asserted. But they are not right, because we should then have a case of the hypothesis of the right angle, which (from the preceding) would leave no place for the hypothesis of the obtuse angle. But neither are they acute. For in that case RX would be greater than AB (3), and therefore still greater than CD. But that this cannot be true is shown as follows. If the quadrilateral $CDXR$ is known to be filled with straight lines cutting off portions from CR and DX equal to each other, this implies a passing from the straight line CD which is less than AB, to RX greater than the same, and so a passing through a certain ST equal to AB. But that this is absurd in our present hypothesis follows from it, because then there would



be one case of the hypothesis of the right angle (4), which would leave no place for the hypothesis of the obtuse angle (from the preceding). Therefore the angles with the joining line *RX* ought to be obtuse.

Then taking on *AC* and *BD* equal portions *AL* and *BK*, we shall show in a similar manner that the angles with the joining line *LK* cannot be acute towards *AB*, because then it would be greater than *AB*, and therefore still greater than *CD*. But from this there ought to be found as above a certain intermediate line between *CD* which is smaller, and *LK* which is larger than *AB*, intermediate I say, and equal to *AB*, which certainly, from what has just been noted, would take away all place for the hypothesis of the obtuse angle. Finally, for this same reason, the angles with the joining line *LK* cannot be right angles. Therefore they will be obtuse. Therefore on the same base *AB*, the perpendiculars being increased or diminished at will, there will always remain the hypothesis of the obtuse angle.

But the same ought to be demonstrated on the assumption of any base. Let there be chosen for base one of the above mentioned perpendiculars, say *BX*. Bisect *AB* and *RX* at the points *M* and *H* and join *M* and *H*. *MH* will be perpendicular to *AB* and *RX* (2). But the angle at *B* is a right angle by hypothesis, and the angle at *X* is obtuse as just proved. Therefore make the right angle *BXP* on the side of *MH*, *XP* will cut *MH* at a certain point *P* situated between the points *M* and *H*, since, on the one hand, the angle *BXH* is obtuse, and, on the other hand, if we join *X* and *M*, the angle *BXM* is acute (I, 17). Then indeed since the quadrilateral *XBMP* contains three right angles from what is already known and one obtuse at the point *P*, because it is exterior with respect to the interior opposite right angle at the point *H* of the triangle *PHX* (I, 16), the side *XP* will be less than the opposite side *BM* (3, Cor. 1). Therefore, taking in *BM* the portion *BF* equal to *XP*, the angles with the joining line *PF* will be equal to each other, and even obtuse, since the angle *BFP* is obtuse on account of the interior opposite angle *FMP* (I, 16). Therefore under any base *BX* the hypothesis of the obtuse angle holds true.



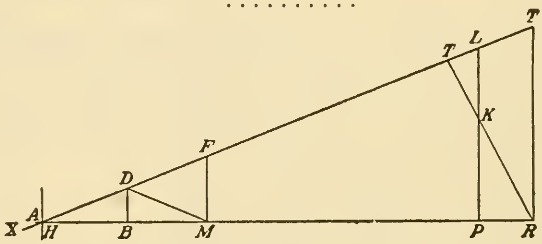
Moreover there stands as above the same hypothesis on the same base BX, however the equal perpendiculars are increased at will or diminished. And so it follows that the hypothesis of the obtuse angle, if it is true in one case, is always in every case the only true hypothesis. Q. E. D.

Proposition VII.—The hypothesis of the acute angle, if it is true in one case, is always in every case the only true hypothesis.

It is proved most easily. For if the hypothesis of the acute angle should permit any case of either the hypothesis of the right angle or of the obtuse angle, already (from the two preceding propositions) no place will be left for the hypothesis of the acute angle, which is absurd. Therefore the hypothesis of the acute angle, if it is true in one case, is always in every case the only true hypothesis.

The following propositions, given here without the proofs, show in part the course pursued by Saccheri.

. . . .



Proposition XI.—Let the straight line AP (as long as you please) cut two straight lines PL and AD, the first in a right angle at P, but the second at A in an angle that is acute towards the side of PL. I say that the straight lines AD and PL (in the hypothesis of the right angle) will at length meet in some point, and indeed at a finite or terminated distance, if produced towards those parts on which they make with the base line AP the two angles less than two right angles.

. . . .

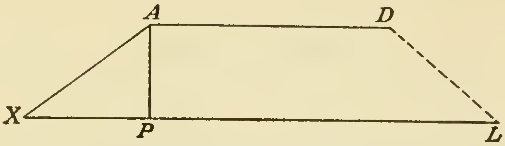
Proposition XII.—Again I say that the straight line AD will meet the line PL somewhere towards these parts (and indeed at a finite or terminated distance) even in the hypothesis of the obtuse angle.

. . . .

Proposition XIII.—If the straight line XA (of any designated length) intersecting two straight lines AD and XL, makes with them on the same side interior angles XAD and AXL less than two right angles, I say that those two (even if neither of those angles is a right angle) will meet each other at length in some point on the side of

those angles, and indeed at a finite or terminated distance, if holds true the hypothesis of the right angle or of the obtuse angle.

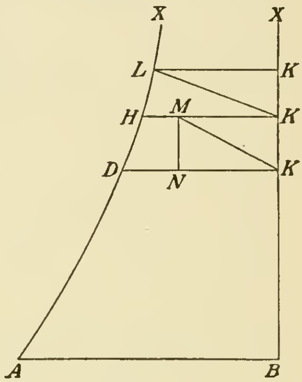
. . . .



Proposition XIV.—The hypothesis of the obtuse angle is absolutely false, because it destroys itself.[1]

. . . .

Proposition XXIII.—If two lines AX and BX lie in the same plane, either they have one common perpendicular (even in the hypothesis of the acute angle), or prolonged, both towards one side or towards the other, unless somewhere one meets the other at a finite distance, they will always more and more nearly approach each other.



. . . .

Proposition XXXIII.—The hypothesis of the acute angle is absolutely false, because repugnant to the nature of a straight line.[2]

[1] [Propositions XII and XIII lead to Euclid's postulate, and so to the hypothesis of the right angle, "even in the hypothesis of the obtuse angle."]

[2] [Up to this point his proofs are clear and logical. But he fails to find his contradiction and falls back on vague illogical reasoning. In the proof of Proposition XXXIII he says, "For then we have two lines which produced must run together into the same line and have at one and the same infinitely distant point a common perpendicular." Then he says he will go into first principles most carefully in order not to omit any objection. Finally in Part II he comes to

Proposition XXXVIII. *The hypothesis of the acute angle is absolutely false, because it destroys itself.*

In the summary at the beginning he says that after the falsity of the hypothesis of the obtuse angle is shown "begins a long battle against the hypothesis of the acute angle," which alone denies the truth of that axiom.]

LOBACHEVSKY

On Non-Euclidean Geometry

(Translated from the French by Professor Henry P. Manning, Brown University, Providence, R. I.)

Nicholas Ivanovich Lobachevsky was born in 1793 and died in 1856. For almost his entire life he was connected with the University of Kasan where he was professor of mathematics and finally rector. He wrote several memoirs and books on the theory of parallels, of which three may be mentioned as the most important: (1) *New Foundations of Geometry*, published first in Russian in 1835–1838. A German translation is given by Engel and Stäckel *Urkunden*, vol. I, pages 67–236. There was an English translation made by Halsted in 1897, and a French translation made in 1901. (2) *Geometrical Investigations on the Theory of Parallels*, written in German and published as a book in Berlin in 1840. This was translated into French by Höuel in 1866, and into English in 1891 by Halsted (Chicago, 1914). (3) *Pangeometry*, published simultaneously in Russian and French in 1855, translated into German in 1858 and again in 1902, and into Italian in 1867. This is more condensed, many proofs being omitted with references to *Geometrical Investigations*, and as it was written near the end of Lobachevsky's life it may be regarded as representing the final development of his ideas. When this was written he had become blind and had to dictate whatever he wrote to his pupils.

Pangeometry

or a Summary of Geometry Founded upon a General and Rigorous Theory of Parallels.[1]

The notions upon which the elementary geometry is founded are not sufficient for a deduction from them of a demonstration of the theorem that the sum of the three angles of a rectilinear triangle is always equal to two right angles, a theorem the truth

---

[1] [*Collection complète des œuvres géométriques de N. I. Lobatcheffsky*, volume II, Kasan, 1886, pages 617–680. This translation has been compared with the Russian edition by Mrs. D. H. Lehmer of Brown University. There are a few differences. Some superfluous words are omitted in the French and obscure passages in the Russian are explained more fully and so made clearer. Apparently the Russian edition was printed first and the French shows some slight revision. Some of these differences will be pointed out below.]

of which no one has doubted to the present time because we meet no contradiction in the consequences which we have deduced from it, and because direct measures of angles of rectilinear triangles agree with this theorem within the limits of error of the most perfect measures.

The insufficiency of the fundamental notions for the demonstration of this theorem has forced geometers to adopt explicitly or implicitly auxiliary suppositions, which, however simple they appear, are no less arbitrary and therefore inadmissible. Thus, for example, one assumes that a circle of infinite radius becomes a straight line, and a sphere of infinite radius a plane, that the angles of a rectilinear triangle always depend only on the ratios of the sides and not on the sides themselves, or, finally, as it is ordinarily done in the elements of geometry, that through a given point of a plane we can draw only a single line parallel to another given line in the plane, while all other lines drawn through the same point and in the same plane ought necessarily to cut the given line if sufficiently prolonged. We understand by the term "line parallel to a given line" a line which, however far it is prolonged in both directions, never cuts the one to which it is parallel. This definition is of itself insufficient, because it does not sufficiently characterize a single straight line. We may say the same thing of most of the definitions given ordinarily in the elements of geometry, for these definitions not only do not indicate the generation of the magnitudes which they define, but they do not even show that these magnitudes can exist. Thus we define the straight line and the plane by one of their properties. We say that straight lines are those which always coincide when they have two points in common, and that a plane is a surface in which a line lies entirely when the line has two points in common with it.

Instead of commencing geometry with the plane and the straight line as we do ordinarily, I have preferred to commence it with the sphere and the circle, whose definitions are not subject to the reproach of being incomplete, since they contain the generation of the magnitudes which they define.

Then I define the plane as the geometrical locus of the intersections of equal spheres described around two fixed points as centers. Finally I define the straight line as the geometrical locus of the intersections of equal circles, all situated in a single plane and described around two fixed points of this plane as centers. If these definitions of the plane and straight line are accepted all the

theory of perpendicular planes and lines can be explained and demonstrated with much simplicity and brevity.[1]

Being given a straight line and a point in a plane, I define as parallel through the given point to the given line, the limiting line between those drawn in the same plane through the same point and prolonged on one side of the perpendicular from the point to the given line, which cut it, and those which do not cut it.[2]

I have published a complete theory of parallels under the title *Geometrical Investigations on the Theory of Parallels*, Berlin, 1840, in the Finck publishing house. In this work I have stated first all the theorems which can be demonstrated without the aid of the theory of parallels. Among these theorems the theorem which gives the ratio of the area of a spherical triangle to the entire area of the sphere upon which it is traced, is particularly remarkable (*Geometrical Investigations*, §27.)[3] If A, B, and C are the angles of a spherical triangle and π represents 2 right angles,the ratio of the area of the triangle to the area of the sphere to which it belongs will be equal to the ratio of

$$\tfrac{1}{2}(A + B + C - \pi)$$

to four right angles.

Then I demonstrate that the sum of the three angles of a rectilinear triangle can never surpass two right angles (§19), and that, if the sum is equal to two right angles in any triangle, it will be so in all (§20). Thus there are only two suppositions possible: Either the sum of the three angles of a rectilinear triangle is always equal to two right angles, the supposition which gives the known geometry, or in every rectilinear triangle this sum is less than two right angles, and this supposition serves as the basis of another geometry, to which I had given the name of *imaginary geometry*, but which it is perhaps more fitting to call *pangeometry* because this name indicates a general geometrical theory which includes the ordinary geometry as a particular case. It follows from the principles adopted in the pangeometry that a perpendicular p let fall from a

---

[1] [He seems to refer to work elsewhere, or perhaps to his teaching. These ideas are not developed further in this book.]

[2] [The Russian adds: That side on which the intersection occurs I call *the side of parallelism*.]

[3] [The *Geometrical Investigations* is the work translated by Halsted. See page 360. In further references to this work only the section number will be given.]

point of a straight line upon one of the parallels makes with the first line two angles of which one is acute. I call this angle the *angle of parallelism* and the side of the first line where it is found,[1] side which is the same for all the points of this line, the *side of parallelism.* I denote this angle by $\Pi(p)$, since it depends upon the length of the perpendicular. In the ordinary geometry we have always $\Pi(p)$ = a right angle for every length of p. In the pangeometry the angle $\Pi(p)$ passes through all values from zero, which corresponds to $p = \infty$, to $\Pi(p)$ = a right angle for $p = 0$ (§23). In order to give the function $\Pi(p)$ a more general analytical value I assume that the value of this function for p negative, case which the original definition does not cover, is fixed by the equation

$$\Pi(p) + \Pi(-p) = \pi.$$

Thus for every angle $A > 0$ and $< \pi$ we can find a line p such that $\Pi(p) = A$, where the line p will be positive if $A < \pi/2$. Reciprocally there exists for every line p an angle A such that $A = \Pi(p)$.

I call *limit circle*[2] the circle whose radius is infinite. It can be traced approximately by constructing in the following manner as many points as we wish. Take a point on an indefinite straight line, call this point *vertex* and the line *axis* of the limit circle, and construct an angle $A > 0$ and $< \pi/2$ with vertex at the vertex of the limit circle and the axis of the limit circle as one of its sides. Then let a be the line which gives $\Pi(a) = A$ and lay off on the second side of the angle from the vertex a length equal to $2a$. The extremity of this length will be found on the limit circle. To continue the tracing of the limit circle on the other side of the axis it will be necessary to repeat the construction on that side. It follows that all the lines parallel to the axis of the limit circle can be taken as axes.

The rotation of the limit circle around one of its axes produces a surface which I call *limit sphere*,[3] surface which is, therefore, the limit which the sphere approaches if the radius increases to infinity. We shall call the axis of rotation, and therefore all the lines parallel to the axis of rotation, *axes of the limit sphere*, and we shall call *diametral plane* every plane which contains one or several axes of the limit sphere. The intersections of the limit sphere by its

---

[1] [Russian: Where the acute angle is found.]

[2] [This is the *oricycle* or *boundry-curve.*]

[3] [*Orisphere* or *boundaiy-surface.*]

diametral planes are limit circles. A part of the surface of the limit sphere bounded by three limit circle arcs will be called a *limit sphere triangle.* The limit circle arcs will be called the *sides*, and the dihedral angles between the planes of these arcs the *angles* of the limit sphere triangle.

Two lines parallel to a third are parallel to each other (§25). It follows that all the axes of a limit circle and of a limit sphere are parallel to one another. If three planes two by two intersect in three parallel lines and if we limit each plane to the part which is between these parallels, the sum of the three dihedral angles which these planes form will be equal to two right angles (§28). It follows from this theorem that the sum of the angles of a limit sphere triangle is always equal to two right angles, and everything that is demonstrated in the ordinary geometry of the proportionality of the sides of rectilinear triangles can therefore be demonstrated in the same manner in the pangeometry of the limit sphere triangles if only we will replace the lines parallel to the sides of the rectilinear triangle by limit circle arcs drawn through the points of one of the sides of the limit sphere triangle and all making the same angle with this side.[1] Thus, for example, if p, q, and r are the sides of a limit sphere right triangle and P, Q, and $\pi/2$ the angles opposite these sides, it is necessary to assume, as for right angled rectilinear right triangles of the ordinary geometry, the equations

$$p = r \sin P = r \cos Q,$$
$$q = r \cos P = r \sin Q,$$
$$P + Q = \frac{\pi}{2}.$$

In the ordinary geometry we demonstrate that the distance between two parallel lines is constant. In pangeometry, on the contrary, the distance p from a point of a line to the parallel line diminishes on the side of parallelism, that is to say, on the side towards which is turned the angle of parallelism $\Pi(p)$.

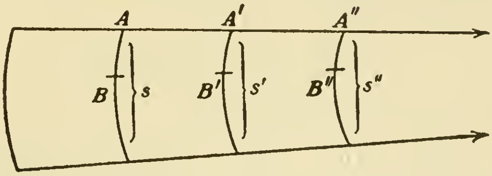
Now let $s, s', s'', \ldots$ be a series of limit circle arcs lying between two parallel lines which serve as axes to all these limit circles, and suppose that the parts of these parallel lines between

[1] [Apparently he would say that two limit circle arcs cutting a third so that corresponding angles are equal would be like the parallels of ordinary geometry. Thus a limit circle arc cutting one side of a limit sphere triangle may be "parallel" to one of the other sides.]

two consecutive arcs are all equal to one another and equal to x. Denote by E the ratio of s to s',[1]

$$\frac{s}{s'} = E,$$

where E is a number greater than unity.[2]



Suppose[3] first that $E = n/m$, m and n being two integer numbers, and divide the arc s into m equal parts. Through the points of division draw lines parallel to the axes of the limit circles. These parallels will divide each of the arcs s', s'', etc., into m parts equal to one another. Let[4] AB be the first part of s, $A'B'$ the first part of s', $A''B''$ the first part of s'' etc., A, A', A'', . . . the points situated upon one of the given parallels, and put $A'B'$ upon AB so that A and A' will coincide and $A'B'$ fall along AB. Repeat this superposition n times. Since by supposition $s/s' = n/m$, it will be necessary that $nA'B' = mAB$, and therefore that the second extremity of $A'B'$ will coincide after the nth superposition with the second extremity of s, which will be divided into n equal parts. s', s'', . . . will also be divided into m equal parts each by the lines parallel to the two given parallels. But if we consider that in making the superposition indicated above, $A'B'$ carries the part of the plane limited by this arc and the two parallels drawn through its extremi-

[1] [The Russian adds: when x is equal to 1.]

[2] [Russian: positive and greater than unity.]

[3] [There are no figures for the pangeometry in the *Œuvres* from which this translation is made. The figures that we are using are taken from the German translation made by Heinrich Liebmann, Leipzig, 1902.]

[4] [Instead of the rest of this paragraph the Russian says: We superimpose the area between the arcs s' and s'' over the area between s and s', and the arc s' on the arc s, and hence s'' on s'. We repeat the arc s'/m. It has to go n times in the arc s. Parallelism of lines makes the arc s''/m go n times in s'. Hence

$$s/s' = s'/s''.$$

And a few lines below: This implies that for every line x, $s' = sE^{-x}$, etc.]

ties, it is clear that at the same time while n times $A'B'$ covers all of the arc s, $nA''B''$ will cover all of the arc s', and so on, because in this case the parallels ought to coincide in all their extent, so that we have

$$nA''B'' = mA'B',$$

or, what is the same thing,

$$\frac{s'}{s''} = \frac{n}{m} = E, \frac{s'}{s''} = E, \text{ etc.},$$

which is what we had to demonstrate.

To demonstrate the same thing in the case where E is an incommensurable number we can employ one of the methods used for similar cases in ordinary geometry. For the sake of brevity I will omit these details. Thus

$$\frac{s}{s'} = \frac{s'}{s''} = \frac{s''}{s'''} = \ldots = E.$$

After this it is not difficult to conclude that

$$s' = sE^{-x},$$

where E is the value of s/s' for x, the distance between the arcs s and s', equal to unity.

It is necessary to remark that this ratio E does not depend on the length of the arc s, and remains the same if the two given parallel lines are moved away from each other or approach each other. The number E, which is necesarily greater than unity, depends only on the unit of length, which is the distance between two successive arcs, and which is entirely arbitrary. The property which we have just demonstrated with respect to the arcs s, s', $s''\ldots$ subsists for the areas P, P', P'', . . ., limited by two successive arcs and the two parallels. We have then

$$P' = PE^{-x}.$$

If we unite n such areas P, P', P'', . . . $P^{(n-1)}$, the sum will be

$$P\frac{1 - E^{-nx}}{1 - E^{-x}}.$$

For $n = \infty$ this expression gives the area of the part of the plane between two parallel lines, limited on one side by the arc s, and unlimited on the side of the parallelism, and the value of this will be

$$\frac{P}{1 - E^{-x}}.$$

If we choose for unit of area the area P which corresponds to an arc s also a unit, and to $x = 1$, we shall have in general for any arc s

$$\frac{Es}{E-1}.$$

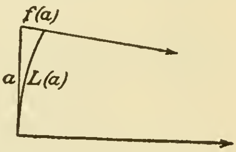
In the ordinary geometry the ratio designated by E is constant and equal to unity. It follows that in the ordinary geometry two parallel lines are everywhere equidistant and that the area of the part of the plane situated between two parallel lines and limited on one side only by a perpendicular common to them is infinite.

Consider for the present a right angled rectilinear triangle in which a, b, and c are the sides, and A, B, and $\pi/2$ the angles opposite these sides. For the angles A and B can be taken the angles of parallelism $\Pi(\alpha)$ and $\Pi(\beta)$ corresponding to lines of positive length α and β. Let us agree also to denote hereafter by a letter with an accent a line whose length corresponds to an angle of parallelism which is the complement to a right angle of the angle of parallelism corresponding to the line whose length is denoted by the same letter without accent, so that we have always

$$\Pi(\alpha) + \Pi(\alpha') = \frac{\pi}{2},$$

$$\Pi(b) + \Pi(b') = \frac{\pi}{2}.$$

Denote[1] by $f(a)$ the part of a parallel to an axis of a limit circle intercepted between the perpendicular to the axis through the vertex of the limit circle and the limit circle itself, if this parallel passes through a point of the perpendicular whose distance from the vertex is a, and let $L(a)$ be the length of the arc from the vertex to this parallel.



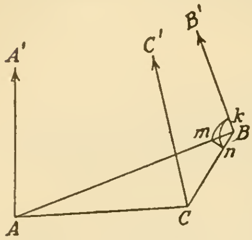
In the ordinary geometry we have

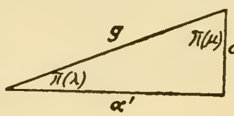
$$f(a) = 0,\ L(a) = a,$$

for every length a.

[1] [The Russian says: Erect a perpendicular a to an axis of a limit circle at the vertex. Through the apex of the perpendicular draw a line parallel to the axis on the side of parallelism. Designate by $f(a)$ the part of the parallel between the perpendicular and the limit circle itself and by $L(a)$ the length of the arc from the vertex to this parallel.]

Draw a perpendicular AA' to the plane of the right angled triangle whose sides have been denoted by a, b, and c, perpendicular through the vertex A of the angle $\Pi(\alpha)$. Pass through this perpendicular two planes, of which one, which we will call the first plane, passes also through the side b, and the other, the second plane, through the side c. Construct in the second plane the line BB' parallel to AA' which passes through the vertex B of the angle $\Pi(\beta)$, and pass a third plane through BB' and the side a of the triangle. This third plane will cut the first in a line CC' parallel to AA'. Conceive now a sphere described from the point B as center with a radius arbitrary, but smaller than a, a sphere which will therefore cut the two sides a and c of the triangle and the line BB' in three points which we will call, the first n, the second m, and the third k. The arcs of great circles, intersections of this sphere by the three planes passing through B, which unite two by two the points n, m, and k, will form a spherical triangle right angled at m, whose sides will be $mn = \Pi(\beta)$, $km = \Pi(c)$, and $kn = \Pi(a)$. The spherical angle knm will be equal to $\Pi(b)$ and the angle kmn will be a right angle. The three lines being parallel to one another, the sum of the three dihedral angles which the parts of the planes $AA'BB'$, $AA'CC'$, and $BB'CC'$ situated between the lines AA', BB', and CC' form with one another will be equal to two right angles.[1] It follows that the third angle of the spherical triangle will be $mkn = \Pi(\alpha')$. We see then that to every right angled rectilinear triangle whose sides are a, b, and c, and the opposite angles $\Pi(\alpha)$, $\Pi(\beta)$, and $\pi/2$ corresponds a right angled spherical triangle whose sides are $\pi(\beta)$, $\Pi(c)$, and $\Pi(a)$, and the opposite angles $\Pi(\alpha')$, $\Pi(b)$, and $\pi/2$. Construct another right angled rectilinear triangle whose sides perpendicular to each other are[2] α' and a, whose hypotenuse is g, and in which $\Pi(\lambda)$ is the angle opposite the side a, and $\Pi(\mu)$ the angle opposite the side α'. Pass from this triangle to the spherical triangle which corresponds in the same manner as the spherical triangle





[1] [The Russian says: The three lines AA', BB', and CC', being parallel to one another, give a sum of dihedral angles equal to π.]

[2] [It would be better to put a before α'.]

kmn corresponds to the triangle *ABC*. The sides of this spherical triangle will then be

$$\Pi(\mu), \Pi(g), \Pi(a),$$

and the opposite angles

$$\Pi(\lambda'), \Pi(\alpha'), \frac{\pi}{2},$$

and it will have its parts equal to the corresponding parts of the spherical triangle *kmn*, for the sides of the latter were

$$\Pi(c), \Pi(\beta), \Pi(a),$$

and the opposite angles

$$\Pi(b), \Pi(\alpha'), \frac{\pi}{2},$$

which shows that these spherical triangles have the hypotenuse and an adjacent angle the same.

It follows that

$$\mu = c,\ g = \beta,\ b = \lambda',$$

and thus the existence of the right angled rectilinear triangle with the sides

$$a,\ b,\ c,$$

and the opposite angles

$$\Pi(\alpha), \Pi(\beta), \frac{\pi}{2},$$

supposes the existence of a right angled rectilinear triangle with the sides

$$a,\ \alpha',\ \beta,$$

and the opposite angles

$$\Pi(b'), \Pi(c), \frac{\pi}{2}.$$

We can express the same thing by saying that if

$$a,\ b,\ c,\ \alpha,\ \beta$$

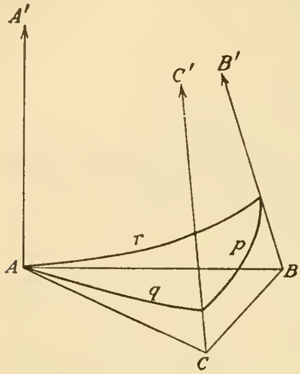
are the parts of a right angled rectilinear triangle,

$$a,\ \alpha',\ \beta,\ b',\ c$$

will be the parts of another right angled rectilinear triangle.[1]

[1] [There seems to be no simple geometrical relation between the two rectilinear triangles nor anything more than an empirical law for deriving the parts of the spherical triangle from those of the rectilinear triangle. There are two ways in which the parts of two right triangles may correspond and there is

If we construct the limit sphere of which the perpendicular AA' to the plane of the given right angled rectilinear triangle is an axis and



the point A the vertex, we shall have a triangle situated upon the limit sphere and produced by its intersections with the planes drawn through the three sides of the given triangle. Denote the three sides of this limit sphere triangle by p, q, and r, p the intersection of the limit sphere by the plane which passes through a, q the intersection by the plane which passes through b, and r the intersection by the plane which passes through c. The angles opposite these sides will be $\Pi(\alpha)$ opposite p, $\Pi(\alpha')$ opposite q, and a right angle opposite r. From the conventions adopted above $q = L(b)$ and $r = L(c)$. The limit sphere will cut the line CC' at a point whose distance from C will be, from the same conventions, $f(b)$. In the same manner we shall have $f(c)$ for the distance from the point of intersection of the limit sphere with the line BB' to the point B.

It is easy to see that we shall have

$$f(b) + f(a) = f(c).$$

---

some confusion here because the correspondence of the two spherical triangles as derived from corresponding rectilinear triangles is not the correspondence in which the parts of one are equal to the corresponding parts of the other.

If we indicate the parts of the first rectilinear triangle by writing

$$a,\ b,\ c,\ \alpha,\ \beta,$$

and the corresponding parts ol the spherical triangle by writing

$$\beta,\ c,\ a,\ \alpha',\ b,$$

we can write for the second rectilinear triangle

$$a,\ \alpha',\ g,\ \lambda,\ \mu,$$

or, substituting for g, λ, and μ their values,

$$a,\ \alpha,'\ \beta,\ b,'\ c,$$

and then for the second spherical triangle

$$c,\ \beta,\ a,\ b,\ \alpha',$$

and we find that the parts of the two spherical triangles are not arranged according to the way in which they are equal.

When he writes the parts of the first spherical triangle for the purpose of comparing the two he changes the order so that the parts of the two are arranged in this way.]

In the triangle whose sides are the limit circle arcs p, q, and r we shall have

$$p = r \sin \Pi(\alpha), \ q = r \cos \Pi(\alpha).$$

Multiplying the first of these two equations by $E^{f(b)}$ we have

$$pE^{f(b)} = r \sin \Pi(\alpha) E^{f(b)}.$$

But

$$pE^{f(b)} = L(a),$$

and therefore

$$L(a) = r \sin \Pi(\alpha) E^{f(b)}.$$

In the same way

$$L(b) = r \sin \Pi(\beta) E^{f(a)}.$$

At the same time $q = r \cos \Pi(\alpha)$, or, what is the same thing, $L(b) = r \cos \Pi(\alpha)$. A comparison of the two values of $L(b)$ gives the equation

$$\cos \Pi(\alpha) = \sin \Pi(\beta) E^{f(a)}. \tag{1}$$

Substituting b' for α and c for β without changing a, which is permitted from what we have demonstrated above, we shall have

$$\cos \Pi(b') = \sin \Pi(c) E^{f(a)},$$

or, since

$$\Pi(b) + \Pi(b') = \frac{\pi}{2},$$

$$\sin \Pi(b) = \sin \Pi(c) E^{f(a)}.$$

In the same way we ought to have

$$\sin \Pi(a) = \sin \Pi(c) E^{f(b)}.$$

Multiply the last equation by $E^{f(a)}$ and substitute $f(c)$ for $f(a) + f(b)$. This will give

$$\sin \Pi(a) E^{f(a)} = \sin \Pi(c) E^{f(c)}.$$

But as in a right angled rectilinear triangle the perpendicular sides can vary while the hypotenuse remains constant, we can put in this equation $a = 0$ without changing c. This will give, since $f(0) = 0$ and $\Pi(0) = \pi/2$,

$$1 = \sin \Pi(c) E^{f(c)},$$

or

$$E^{f(c)} = \frac{1}{\sin \Pi(c)}$$

for every line c.

Now take equation (1)

$$\cos \Pi(\alpha) = \sin \Pi(\beta) E^{f(a)}$$

and substitute $1/\sin \Pi(a)$ for $E^{f(a)}$. It will take the following form

$$\cos \Pi(\alpha) \sin \Pi(a) = \sin \Pi(\beta). \tag{2}$$

Changing α and β to b' and c without changing a we find

$$\sin \Pi(b) \sin \Pi(a) = \sin \Pi(c).$$

Equation (2) with a change of letters gives

$$\cos \Pi(\beta) \sin \Pi(b) = \sin \Pi(\alpha).$$

If in this equation we change β, b, and α to c, α', and b' we shall get

$$\cos \Pi(c) \cos \Pi(\alpha) = \cos \Pi(b). \tag{3}$$

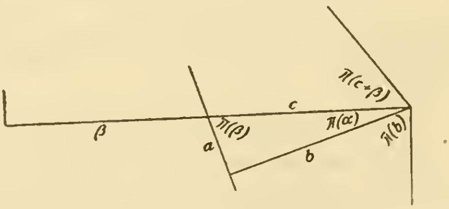
In the same way we shall have

$$\cos \Pi(c) \cos \Pi(\beta) = \cos \Pi(a) \tag{4}$$

. . . .

It follows[1] from what precedes that spherical trigonometry remains the same, whether we adopt the supposition that the sum of the three angles of a rectilinear triangle is always equal to two right angles, or adopt the supposition that this sum is always less than two right angles, which is very remarkable and does not hold for rectilinear trigonometry.

Before demonstrating the equations which express in pangeometry the relations between the sides and angles of any rectilinear triangle we shall seek for every line x the form of the function which we have denoted hitherto by $\Pi(x)$.

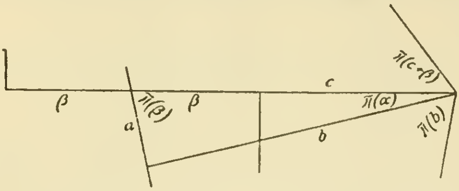


Consider[2] for this purpose a right angled rectilinear triangle whose sides are a, b, c, and the opposite angles $\Pi(\alpha)$, $\Pi(\beta)$, $\pi/2$. Prolong c beyond the vertex of the angle $\Pi(\beta)$ and make the

[1] [In the part omitted the ordinary equations of spherical trigonometry are derived from the preceding equations.]

[2] [We have combined and changed a little the figures given here by Liebmann.]

prolongation equal to β. The perpendicular to β erected at the extremity of this line and on the side opposite to that of the angle $\Pi(\beta)$ will be parallel to a and its prolongation beyond the vertex of $\Pi(\beta)$. Draw also through the vertex of $\Pi(\alpha)$ a line parallel



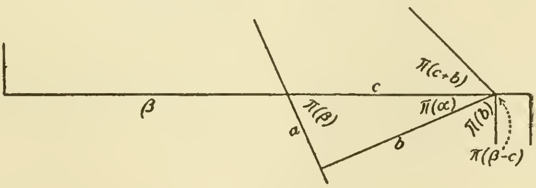
to this same prolongation of a. The angle which this line will make with c will be $\Pi(c + \beta)$ and the angle which it will make with b will be $\Pi(b)$, and we shall have the equation

$$\Pi(b) = \Pi(c + \beta) + \Pi(\alpha). \tag{Π}$$

If we take the length β from the vertex of the angle $\Pi(\beta)$ on the side c itself and erect at its extremity a perpendicular to β on the side of the angle $\Pi(\beta)$, this line will be parallel to the prolongation of a beyond the vertex of the right angle. Draw through the vertex of the angle $\Pi(\alpha)$ a line parallel to this last perpendicular, which will therefore also be parallel to the second prolongation of a. The angle of this parallel with c will be in all cases $\Pi(c - \beta)$ and the angle which it makes with b will be $\Pi(b)$. Therefore

$$\Pi(b) = \Pi(c - \beta) - \Pi(\alpha). \tag{Π'}$$

It is easy to convince ourselves that this equation is true not only if $c > \beta$, but also if $c = \beta$ and if $c < \beta$. In fact, if $c = \beta$ we have, on the one hand, $\Pi(c - \beta) = \Pi(0) = \pi/2$, and, on the other hand, the perpendicular to c drawn through the vertex of the angle $\Pi(\alpha)$ becomes parallel to a, whence it follows that $\Pi(b) = \frac{\pi}{2} - \Pi(\alpha)$, which agrees with our equation.



If $c < \beta$ the extremity of the line β will fall beyond the vertex of the angle $\Pi(\alpha)$ at a distance equal to $\beta - c$. The perpendicular to β at this extremity of β will be parallel to a and to the line

through the vertex of the angle $\Pi(\alpha)$ parallel to a, whence it follows that the two adjacent angles which this parallel makes with c will be, the acute equal to $\Pi(\beta - c)$, the obtuse equal to $\Pi(\alpha) + \Pi(b)$. But the sum of two adjacent angles is always equal to two right angles. Thus

$$\Pi(\beta - c) + \Pi(\alpha) + \Pi(b) = \pi,$$

or

$$\Pi(b) = \pi - \Pi(\beta - c) - \Pi(\alpha).$$

But from the definition of the function $\Pi(x)$

$$\pi - \Pi(\beta - c) = \Pi(c - \beta),$$

which gives

$$\Pi(b) = \Pi(c - \beta) - \Pi(\alpha),$$

that is to say, the equation found above, which is thus demonstrated for all cases.

The two equations (Π) and (Π') can be replaced by the following two

$$\Pi(b) = \tfrac{1}{2}\Pi(c + \beta) + \tfrac{1}{2}\Pi(c - \beta)$$
$$\Pi(\alpha) = \tfrac{1}{2}\Pi(c - \beta) - \tfrac{1}{2}\Pi(c + \beta).$$

But equation (3) gives us

$$\cos \Pi(c) = \cos \Pi(b)/\cos \Pi(\alpha),$$

and in substituting in this equation in place of $\Pi(b)$ and $\Pi(\alpha)$ their values we get

$$\cos \Pi(c) = \frac{\cos [\tfrac{1}{2}\Pi(c + \beta) + \tfrac{1}{2}\Pi(c - \beta)]}{\cos [\tfrac{1}{2}\Pi(c - \beta) - \tfrac{1}{2}\Pi(c + \beta)]}.$$

From this equation we deduce the following

$$\tan^2 \tfrac{1}{2}\Pi(c) = \tan \tfrac{1}{2}\Pi(c - \beta) \tan \tfrac{1}{2}\Pi(c + \beta).$$

As the lines c and β can vary independently of each other in a right angled rectilinear triangle, we can put successively in the last equation $c = \beta$, $c = 2\beta, \ldots c = n\beta$, and we conclude from the equations thus deduced that in general for every line c and for every positive integer n

$$\tan^n \tfrac{1}{2}\Pi(c) = \tan \tfrac{1}{2}\Pi(nc).$$

It is easy to demonstrate the truth of this equation for n negative or fractional, whence it follows that in choosing the unit of length so that we have $\tan \tfrac{1}{2}\Pi(1) = e^{-1}$, where e is the base of Naperian logarithms, we shall have for every line x

$$\tan \tfrac{1}{2}\Pi(x) = e^{-x}.$$

This expression gives $\Pi(x) = \pi/2$ for $x = 0$, $\Pi(x) = 0$ for $x = \infty$, and $\Pi(x) = \pi$ for $x = -\infty$, agreeing with what we have adopted and demonstrated above.

BOLYAI

On Non-Euclidean Geometry

(Translated from the Latin by Professor Henry P. Manning, Brown University, Providence, R. I.)

János Bolyai (1802–1860) was the son of Farkas Bolyai, a fellow student of Gauss's at Göttingen. Farkas wrote to Gauss in 1816 that his son, then a boy of fourteen, had already a good knowledge of the calculus and its applications to mechanics. János went to the engineering school at Vienna at the age of sixteen and entered the army at the age of twenty-one. About 1825 or 1826 he worked out his theory of parallels and published it in 1832 as an appendix to the first part of a work by his father, the book having the imprimatur of 1829. It was in Latin, but was later translated into French (1867), Italian (1868), German (1872), and English (by Halsted, 1891). The title of the father's work is *Tentamen Juventutem Studiosam in Elementa Matheseos Purae...introducendi.* It appeared in two parts at Maros-Vasarhely, in 1832, 1833. It is the appendix to the first part that is here translated.

APPENDIX[1]

exhibiting the absolutely true science of space,[2] independent of Axiom XI[3] of Euclid (not to be decided a priori), with the geometrical quadrature of a circle in the case of its falsity.

Explanation of signs[4]

$\overline{AB}$[5] denotes the complex of all the points on a line with the points A and B.

---

[1] [Ioannis Bolyai de Bolya, *Appendix*, editio nova, published by the Hungarian Academy of Science, Budapest, 1902, in honor of the centennial aniversary of the author's birth. This was published, as originally, along with the *Tentamen* of his father, and also separately.]

[2] [We may remark that "absolutely true science of space" is a different thing from "absolute science of space," or "absolute geometry," which are the terms often used and seem to have been used sometimes by Bolyai himself.]

[3] [Euclid's axiom of parallels which the best authorities now call "Postulate V."]

[4] [In the edition of the Hungarian Academy points are denoted by small letters in German type. We shall use capital italic letters as is customary in modern textbooks in geometry. Also in that edition parentheses are used much more frequently than with us, often a clause that is an essential part of a sentence is inclosed in parentheses. We shall omit most of these parentheses.]

[5] [Two or more letters without any mark over them denote a limited figure, while a figure is unlimited in a part denoted by a letter with a mark over it.

$A\overline{B}$ denotes that half of the line AB cut at A which contains the point B.

$\overline{ABC}$ denotes the complex of all the points which are in the same plane with the points A, B, and C (these not lying in the same straight line).

$AB\overline{C}$ denotes the half of the plane ABC cut apart by $\overline{AB}$ that contains the point C.

ABC denotes the smaller of the portions into which $\overline{ABC}$ is divided by the complex of $B\overline{A}$ and $B\overline{C}$, or the angle whose sides[1] are $B\overline{A}$ and $B\overline{C}$.

$ABCD$[2] denotes (if D is in ABC and $B\overline{A}$ and $C\overline{D}$ do not cut each other) the portion of ABC enclosed by $B\overline{A}$, BC, and $C\overline{D}$. But $BACD$ is the portion of the plane $\overline{ABC}$ between $\overline{AB}$ and $\overline{CD}$.

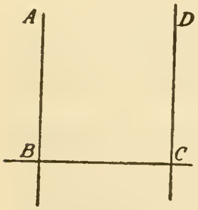
R denotes right angle.

$AB \stackrel{\frown}{=} CD$[3] denotes $CAB = ACD$.

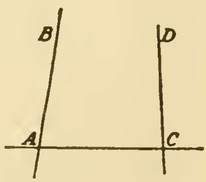
AB (not given in the list) denotes the segment from A to B of the line $\overline{AB}$; and so, in the definition of $ABCD$ below, BC denotes a segment. But we may note that ABC denotes an angle and not a triangle. When the author wishes to name a triangle he says "triangle ABC" or inserts the sign "Δ" (see, for example, §13). We may note also that an angle with him is a portion of a plane, and that two angles having a side in common but lying in different planes will form a dihedral angle. See, for example, §7.]

[1] [He calls them *legs*.]

[2] [If two lines in a plane are cut by a third, we can say that the half-lines on one side of this third line lie in one direction along the two given lines, and the half-lines on the other side lie in the other direction. Now in the first of the two definitions given here we read the two pairs of points taken on the two lines in opposite directions, and in the second definition we read them in the same direction. We have an illustration of the first definition in $MACN$ at the beginning of §2, and an illustration of the second in $BNCP$ at the beginning of §7.]



[3] In the relation represented by this sign AB and CD are two lines in a plane cut by a third at A and C, with the points A and B taken on one line and the points C and D on the other in the same direction. This sign is used when $\overline{AB}$ and $\overline{CD}$ intersect as well as when they do not intersect. See, for example, §5, where $EC \stackrel{\frown}{=} BC$.

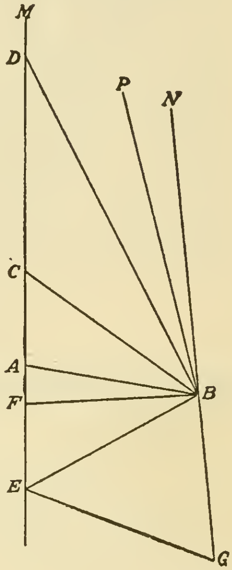


$\equiv$ denotes congruent.[1]
$x \rightarrow a$[2] denotes x tends to a as limit.
$\bigcirc r$ denotes circumference of circle of radius r.
$\odot r$ denotes area of circle of radius r.

§1

Given $\overline{AM}$, if $\overline{BN}$, lying in the same plane with it, does not cut it, but every half-line $\overline{BP}$ in ABN[3] does cut it, let this be denoted by

$$BN|||AM.$$[4]



It is evident that there is *given* such a $\overline{BN}$ and indeed from any point B outside of $\overline{AM}$ *only one*, and that

$$BAM + ABN \text{ is not} > 2R;$$

for when BC[5] is moved around B until

$$BAM + ABC = 2R,$$

at some point $\overline{BC}$ *first* does not cut $\overline{AM}$, and then $BC|||AM$. And it is evident that $BN|||EM$, wherever E may be on $\overline{AM}$ (supposing in all these cases that $AM > AE$).[6]

And if, with the point C on $\overline{AM}$ going off to infinity, we always have $CD = CB$, always we shall have

$$CDB = CBD < NBC.$$

But $NBC \rightarrow 0$. Therefore $ADB \rightarrow 0$.

---

[1] [Footnote to the original] Let it be permitted by this sign, by which Gauss, supreme in geometry, has indicated congruent numbers, to denote also geometrical congruence, since no resulting ambiguity is to be feared.

[2] [Bolyai uses the sign "∽."]

[3] [In the angle ABN, the words "in ABN" are in parentheses in the original. See page 375, footnote 4.]

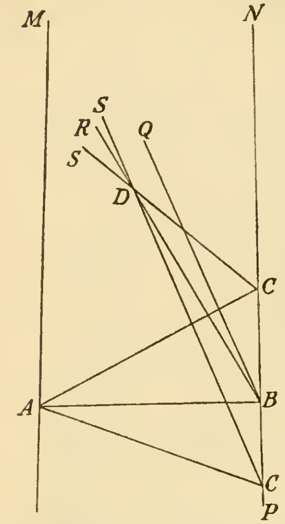
[4] [This is a relation of half-lines, but in writing it the author always leaves out the mark over the second letter.]

[5] [Here he speaks of the segment BC because he thinks of the point C moving along $\overline{AM}$, but for the limiting position he writes $\overline{BC}$.]

[6] [This seems to mean that E is not to lie beyond M on $\overline{AM}$, or that M has been taken far enough out on this half-line to be beyond where we wish to take E.]

§2

If BN|||AM, we shall have also CN|||AM.[1]



For let D be somewhere in $MACN$. If C lies on $\overline{BN}$, $\overline{BD}$ will cut $\overline{AM}$ because $BN|||AM$, and so also $\overline{CD}$ will cut $\overline{AM}$. But if C is in $\overline{BP}$, let $BQ|||CD$. $\overline{BQ}$ falls in ABN (§1)[2] and cuts $\overline{AM}$. and so $\overline{CD}$ cuts $\overline{AM}$. Therefore $\overline{CD}$ cuts $\overline{AM}$ in both cases. But $\overline{CN}$ does not cut $\overline{AM}$. Therefore always $CN|||AM$.

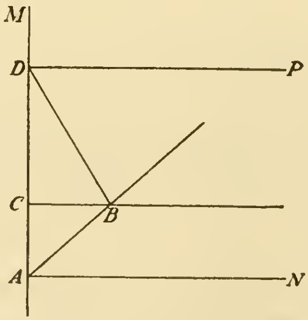
§3

If BR and CS are both |||AM and C is not in $\overline{BR}$, then $\overline{BR}$ and $\overline{CS}$ do not intersect each other.

For if $\overline{BR}$ and $\overline{CS}$ had a point D in common, DR and DS would at the same time $|||AM$ (§2), and $\overline{DS}$ would fall on $\overline{DR}$ (§1) and C on $\overline{BR}$, contrary to hypothesis.

§4

If MAN > MAB, for every point B in $\overline{AB}$ is given a point C in $\overline{AM}$ such that BCM = NAM.



---

[1] [It is left to the reader to see from the figure that C is a point of BN. The author often omits details in this way when they are shown in the figure.]

[2] [BN does not cut $\overline{CD}$, nor will it do so, even if it rotate towards $\overline{BA}$, as long as the two lines intersect below B.]

For there is given a $BDM > NAM$ (§1)[1], and also an $MDP = MAN$, and B falls in $NADP$. If therefore we move NAM along AM until $A\overline{N}$ comes to $D\overline{P}$, sometime $A\overline{N}$ will have passed through B and there will be a $BCM = NAM$.

§5

If $BN|||AM$ (p. 377) *there is a point* F *in* $A\overline{M}$ *such that* $FM \bumpeq BN$.

For there is a $BCM > CBN$ (§1), and if $CE = CB$, and so $EC \bumpeq BC$, it is evident that $BEM < EBN$. Let P traverse EC, the angle BPM always called u and the angle PBN always called[2] v. It is evident that u is at first less than the corresponding value of v, but afterwards greater. But u increases from BEM to BCM *continuously*, since there is *no* angle $> BEM$ and $< BCM$ to which u is not at some time equal (§4). Likewise v decreases from EBN to CBN continuously. And so there is given on EC a point F such that $BFM = FBN$.

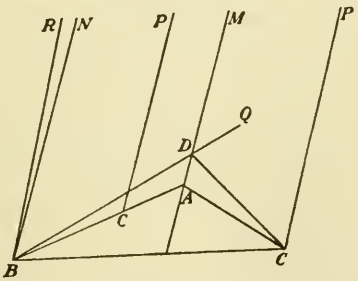
§6

If $BN|||AM$ *and* E *is anywhere in* $\overline{AM}$ *and* G *in* $\overline{BN}$, *then* $GN|||EM$ *and* $EM|||GN$.

For $BN|||EM$ (§1), and hence $GN|||EM$(§2). If then $FM \bumpeq BN$ (§5), then $MFBN \equiv NBFM$, and so, since $BN|||FM$, also $FM|||BN$, and by what precedes $EM|||GN$.

§7

If BN *and* CP *are both* $|||AM$, *and* C *is not in* $\overline{BN}$, *then* $BN|||CP$.



[1] [If we let D move off on $\overline{AM}$, the angle ADB will approach zero, and therefore at some time become less than the supplement of NAM. Then if M is taken beyond this position of D, we shall have BDM the supplement of ADB greater than NAM.]

[2] [To show this in the figure on p. 377, P ought to be put where A is.]

For $B\bar{N}$ and $C\bar{P}$ do not cut each other (§3). Moreover, AM, BN, and CP are in a plane or not. And in the first case AM lies in $BNCP$[1] or not.

If AM, BN, and CP are in a plane and AM falls in $BNCP$, then any $B\bar{Q}$ in NBC cuts $\overline{AM}$ in a point D because $BN|||AM$; and then since $DM|||CP$ (§6) it is evident that $D\bar{Q}$ cuts $C\bar{P}$, and so $BN|||CP$.

But if BN and CP lie on the same side of AM, then one of them, for example CP, will fall *between* the other two, $\overline{BN}$ and $\overline{AM}$, and any $B\bar{Q}$ in NBA will cut $A\bar{M}$, and so also $\overline{CP}$. Therefore $BN|||CP$.

If MAB and MAC form an *angle*,[2] then CBN has in common with ABN only $B\bar{N}$, but $A\bar{M}$ in ABN with $B\bar{N}$, and so NBC with $A\bar{M}$, have in common nothing. But a $BC\bar{D}$ drawn through any $B\bar{D}$ in NBA will cut $A\bar{M}$ because $B\bar{D}$ cuts $A\bar{M}$, BN being$|||AM$. Therefore if $BC\bar{D}$ is moved about BC[3] until *first* it leaves $A\bar{M}$, at last $BC\bar{D}$ will fall in $BC\bar{N}$. For the same reason it will fall in $BC\bar{P}$. Therefore BN falls in $BC\bar{P}$. Then if $BR|||CP$, because also $AM|||CP$, BR will fall in BAM for the same reason, and in BCP because $BR|||CP$. And so $B\bar{R}$ is common to MAB and PCB, and is therefore $B\bar{N}$ itself.[4] Therefore $BN|||CP$.

If therefore $CP|||AM$ and B is outside of $\overline{CAM}$, then the intersection of BAM and BCP, that is, $B\bar{N}$, is $|||$ both to AM and to CP.[5]

---

[1] [Notice that the points B and N are taken on one line and the points C and P on the other in the same direction Thus "in $BNCP$" means in the entire strip between $\overline{BN}$ and $\overline{CP}$, and not simply in that portion of this strip which is above BC. In this paragraph we should take the CP that is to the right in Figure 58 and regard the entire figure as lying in one plane. In the second paragraph we take the CP at the left, while in the third paragraph we take the CP again at the right, but this time the three lines not in one plane.]

[2] [A dihedral angle. See page 375, end of footnote 5.]

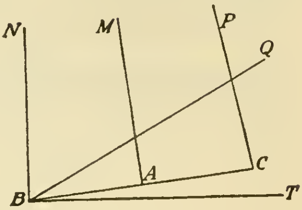
[3] [The point D moving off indefinitely on $\overline{AM}$. This brings $B\bar{D}$ into coincidence with $B\bar{N}$ and $C\bar{D}$ with $C\bar{P}$.]

[4] [The argument is this: BN, which is $|||$ AM, lies in BCP as well as in BAM, and is therefore their intersection. But in the same way we can say that BR, which is $|||$ CP lies in BAM as well as in BCP and is also their intersection.]

[5] [Footnote to the original] If the third case had been taken first, the other two could have been solved as in §10 more briefly and more elegantly (from edition I, volume I, Errata of the Appendix).

§8

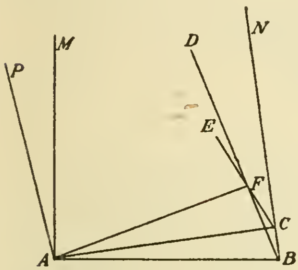
If $BN|||$ *and* $\simeq CP$ *(or more briefly* $BN||| \simeq CP$*), and if* AM *in* $NBCP$ *bisects* BC *at right angles, then* $BN|||AM$.



For if $\overline{BN}$ should cut $\overline{AM}$, also $\overline{CP}$ would cut $\overline{AM}$ at the same point, since $MABN \equiv MACP$, which would be common to $\overline{BN}$ and $\overline{CP}$, although $BN|||CP$. But if $\overline{BQ}$ in CBN cuts $\overline{CP}$ then also $\overline{BQ}$ cuts $\overline{AM}$. Therefore $BN|||AM$.

§9[1]

If $BN\ |||AM$, *if* $MAP \perp MAB$, *and if the angle which* NBD *makes with* NBA *on the side of* $MABN$ *where* MAP *is, is* $< R$, *then* MAP *and* NBD *cut each other.*



For let $BAM = R$, let $AC \perp BN$ (whether B falls at C or not), and let $CE \perp BN$ in NBD. ACE will be $< R$ by hypothesis and $AF \perp CE$ will fall in ACE. Let $\overline{AP}$ be the intersection of $\overline{ABF}$ and $\overline{AMP}$ (these having the point A in common). Then $BAP = BAM = R$ (since $BAM \perp MAP$). If, finally, $\overline{ABF}$ be put upon $\overline{ABM}$, A and B remaining fixed,[2] $\overline{AP}$ will fall on $\overline{AM}$, and since $AC \perp BN$ and $AF < AC$, it is evident that AF will end on this side of $\overline{BN}$ and so BF will fall in ABN. But $\overline{BF}$ will cut $\overline{AP}$ in this position because $BN|||AM$, and so also in their *first* positions $\overline{AP}$ and $\overline{BF}$ will cut each other. The point of intersection is a point common to $\overline{MAP}$ and $\overline{NBD}$, and so $\overline{MAP}$ and $\overline{NBD}$ cut each other.

---

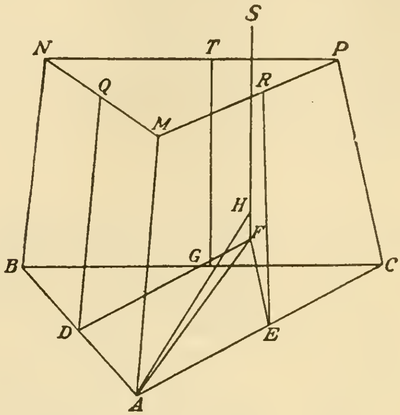
[1] [It will be noticed particularly in this section that he sometimes uses a letter in naming a line or plane and later defines the letter more specifically. Thus at the beginning $\overline{BN}$ and $\overline{AM}$ are arbitrarily given, $BN\ |||\ AM$, but A and B are not both arbitrarily given on these lines, for later they are taken so that $AB \perp AM$. Then he speaks of the half-plane $\overline{MAP}$ although later he takes $\overline{AP}$ as the intersection of this half-plane and another, $\overline{ABF}$, and $\overline{NBD}$ is mentioned twice before he draws $AF \perp CE$, thus determining $\overline{BD}$ apparently as drawn through F.]

[2] [We should say, If $\overline{ABF}$ is revolved on AB so as to fall upon $\overline{ABM}$.]

Then it follows easily that $M\overline{AP}$ and $N\overline{BD}$ mutually intersect, if the sum of the interior angles which they make with $MABN$ is $< 2R$.[1]

§10

If BN and CP are both $|||\ \simeq AM$, *then also* $BN|||\simeq CP$.[2]



For MAB and MAC either make an *angle* or are in a plane.

If the former let $\overline{QDF}$ bisect at right angles the line AB. DQ will be $\perp AB$, and so $DQ|||AM$ (§8). Likewise, if $\overline{ERS}$ bisect AC at right angles, then $ER|||AM$, and therefore $DQ|||ER$ (§7). Easily (through §9) it follows that $\overline{QDF}$ and $\overline{ERS}$ mutually intersect,[3] and the intersection $\overline{FS}$ is $|||DQ$ (§7), and since $BN|||DQ$

[1] [If he means when neither plane is $\perp MABN$, then at least we can say that he does not prove it. See footnote 3 below.]

[2] [We should remember that this theorem has already been proved so far as the first sign, $|||$, is concerned (§7). It is only the equality of angles represented by the sign $\simeq$ that has to be proved here.]

[3] [The theorem of §9 as proved does not seem to apply here directly, for neither of these two planes is perpendicular to the plane of $D\overline{Q}$ and $E\overline{R}$, but a proof can easily be given analogous to the proof of §9. $\overline{QDF}$ is perpendicular to $\overline{ABC}$, and if $\overline{DH}$ is their intersection (not drawn in the figure) the perpendicular from E to $\overline{DH}$, which will fall upon $D\overline{H}$ since it cannot intersect the perpendicular $D\overline{A}$ (Euclid I, 28), will be shorter than any other line from E to $\overline{QDF}$, and so shorter than any line from E to $\overline{DQ}$. If then we revolve HDE around DE until it falls into the plane of $QDER$, $D\overline{H}$ will fall within the angle QDE and will intersect $E\overline{R}$. In the same way the intersection of $RE\overline{S}$ with $\overline{ABC}$, if carried in this revolution into the plane of $QDER$, will fall in the angle

is also $FS|||BN$. Then for every point of $\overline{FS}$ is $FB = FA = FC$,[1] and FS falls in the plane TGF bisecting the line BC at right angles. But (by §7), since $FS|||BN$, also $GT|||BN$. In the same way we prove $GT|||CP$. Moreover GT bisects the line BC at right angles, and so $TGBN \equiv TGCP$ (§1)[2] and $BN||| \simeq CP$.

If BN, AM, and CP are in a plane let FS, falling *outside* of this plane, be $||| \simeq AM$. Then (by the preceding) $FS||| \simeq$ both BN and CP, and so $BN||| \simeq CP$.

§11

Let the complex of the point A and of *all* the points of which any one B is such that f $BN|||AM$, then $BN \simeq AM$, be called F, and let the section of F by any plane containing the line AM be called L.

In any line which is $|||$ AM F has one point and only one, and it is evident that L is divided by AM into two congruent parts. Let $\overline{AM}$ be called the *axis* of L. It is evident also that in any plane containing AM there will be one L *with axis* $\overline{AM}$. Any such L will be called the L *of the axis* $\overline{AM}$, in the plane considered. It is evident that if L is revolved about AM the F will be described of which $\overline{AM}$ is called the *axis*, and conversely the F may be *attributed to the axis* $\overline{AM}$.

§12

If B *is anywhere in the* L *of* $\overline{AM}$ *and* $BN||| \simeq AM$ (§11), *then the* L *of* $\overline{AM}$ *and the* L *of* $\overline{BN}$ *coincide.*

For let the L of $\overline{BN}$ be called for distinction l, and let C be anywhere in l, and $CP||| \simeq BN$ (§11). Then, since also $BN||| \simeq AM$ will CP be $||| \simeq AM$ (§10), and so C will fall also in L. And if C is anywhere in L and $CP||| \simeq AM$, then $CP||| \simeq BN$ (§10), and C will fall also in l (§11). Therefore L and l are the same, and any $\overline{BN}$ is also axis of L and is $\simeq$ among all the axes of L.

The same in the same manner is evident of F.

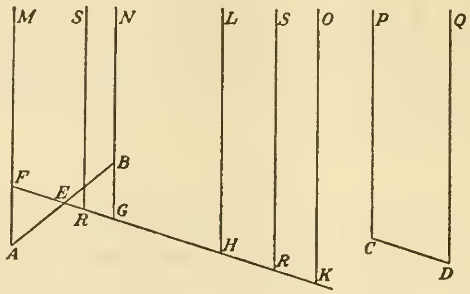
RED and will intersect $\overline{DQ}$. In the plane of $QDER$ those two half-lines must intersect each other, and so before the revolution they must have intersected each other, and their intersection was a point common to $Q\overline{DF}$ and $R\overline{ES}$.]

[1] [Every point of FS is equidistant from A, B, and C.]

[2] [If $TGCP$ is placed upon $TGBN$ (revolved about TG) so that GC falls on GB, CP and BN drawn from the same point $|||$ GT must coincide by §1.]

§13

If $BN|||AM$ and $CP|||DQ$, and $BAM + ABN = 2R$, then also $DCP + CDQ = 2R$.



Let $EA = EB$ and $EFM = DCP$ (§4), then, since

$$BAM + ABN = 2R = ABN + ABG,$$

will

$$EBG = EAF,$$

and if also $BG = AF$,

$$\Delta EBG = \Delta EAF,$$

and

$$BEG = AEF$$

and G falls in $\overline{FE}$. Then is $GFM + FGN = 2R$ (because $EGB = EFA$). Also $GN|||FM$ (§6), and so if $MFRS \equiv PCDQ$, then $RS |||GN$ (§7) and R falls in or outside of FG (if CD is not $= FG$, where the thing is evident).

I. In the first case FRS is not $> 2R - RFM = FGN$ because $RS|||FM$. But since $RS|||GN$ also FRS is not $< FGN$, and so $FRS = FGN$, and

$$RFM + FRS = GFM + FGN = 2R.$$

Therefore $DCP + CDQ = 2R$.

II. If R falls outside of FG, then $NGR = MFR$, and we can let $MFGN = NGHL = LHKO$, and so on until FK first becomes $=$ or $> FR$. This $KO|||HL|||FM$ (§7). If K falls on R then KO falls on RS (§1), and so

$$RFM + FRS = KFM + FKO = KFM + FGN = 2R;$$

but if R falls in HK, then (from I)

$$RHL + HRS = 2R = RFM + FRS = DCP + CDQ.$$

§14

If $BN|||AM$ and $CP|||DQ$ and $BAM + ABN < 2R$, then also $DCP + CDQ < 2R$.

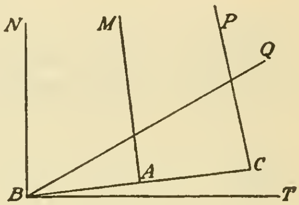
For if $DCP + CDQ$ is not $<$, and so (§1) is $= 2R$, then (by §13) also $BAM + ABN = 2R$, contrary to hypothesis.

§15

Weighing carefully §§13 and 14, *let the system of geometry resting on the hypothesis of the truth of Euclid's Axiom XI be called Σ, and let that one built on the contrary hypothesis be S. All things which are not expressly declared be in Σ or S are to be understood to be announced absolutely, that is, to be true whether Σ or S is true.*

§16

If AM is the axis of any L, then L in Σ is a straight line $\perp AM$.



For at any point B of L let the axis be BN. In Σ

$$BAM + ABN = 2BAM = 2R,$$

and so $BAM = R$. And if C is any point in $\overline{AB}$ and $CP|||AM$, then (by §13) $CP \simeq AM$ and C is in L (§11).

But in S no three points A, B, C, of L or F are in a straight line.

For one of the axes AM, BN, or CP (for example AM) falls between the other two, and then (§14) both BAM and $CAM < R$.

§17

L is also in S a line and F a surface.

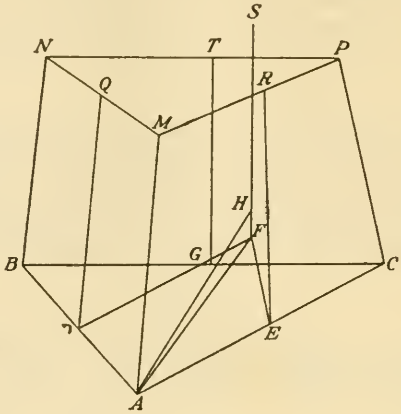
For (from §11) any plane perpendicular to the axis $\overline{AM}$ through any point of F will cut F in the circumference of a circle whose plane is not perpendicular to any other axis $\overline{BN}$ (§14). Let F revolve about BN. Every point of F will remain in F (§12) and the section of F by a plane not perpendicular to $\overline{BN}$ will describe a surface And F (by §12), whatever are the points A and B in it, can be made congruent to *itself* in such a way that A will fall at B. Therefore F is a *uniform surface.*

Hence it is evident (§11 and §12) that L is a *uniform line.*[1]

[1] [Footnote to the original] It is not necessary to restrict the demonstration to S, since the statement may easily be made so as to hold absolutely (for S and Σ). (Edition I, volume I, Errata of the Appendix.)

§18

The section of any plane through a point A of F, oblique to the axis AM, with F in S is the circumference of a circle.



For let *A*, *B*, and *C* be three points of this section, and *BN* and *CP* axes. *AMBN* and *AMCP* will make an angle. For otherwise the plane determined by *A*, *B*, and *C* (§16) will contain *AM*, contrary to hypothesis. Therefore the planes bisecting at right angles *AB* and *AC* will intersect each other (§10) in an axis *FS* of *F*, and $FB = FA = FC$. Let *AH* be $\perp$ *FS*, and revolve *FAH* about *FS*. *A* will describe a circumference of radius *HA* going through *B* and *C*, lying *at the same time* in *F* and in $\overline{ABC}$, nor will *F* and $\overline{ABC}$ have anything in common except ○*HA*.

It is evident also that ○*HA* will be described by the extremity of the portion *FA* of the line *L* (like a radius) rotating in *F* about *F*.[1]

§19

The perpendicular BT to the axis BN of L, falling in the plane of L, is in S tangent to L. [See the figure on p. 385.]

For *L* has no point in $\overline{BT}$ except *B* (§14), but if *BQ* falls in *TBN*, then the center of the plane section through *BQ* perpendicular to *TBN* with the *F* of $\overline{BN}$ is manifestly located in $\overline{BQ}$,[2] and if

[1] [In the surface *F* about the point *F*. In the original there is not this confusion because points are denoted by small German letters. The point *F* is here taken on *FS* so that $AM \simeq FS$.]

[2] [Apparently because the entire figure is symmetrical with respect to the plane *TBN*, and therefore the section is symmetrical with respect to the line *BQ*.]

BQ is the diameter it is evident that $\overline{BQ}$ cuts the L of $\overline{BN}$ in Q.

§20

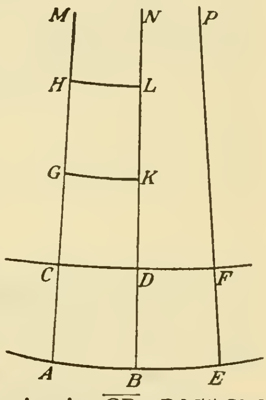
Through any two points of F a line L is determined (§11 and §18), and, since from §16 and §19 L is perpendicular to all its axes, *any L-angle in F is equal to the angle of the planes through its sides perpendicular to F.*

§21

Two L-lines $A\overline{P}$ and $B\overline{D}$ in the same F making with a third L-line AB a sum of interior-angles $< 2R$ intersect each other. [By $\overline{AP}$ in F is meant the L drawn through A and P, and by $A\overline{P}$ that half of it beginning at A in which P falls. See the second figure, p. 381.]

For if AM and BN are axes of F, then $AM\overline{P}$ and $BN\overline{D}$ cut each other (§9),[1] and F cuts their intersection (§7 and §11), and so $A\overline{P}$ and $B\overline{D}$ mutually intersect.

It is evident from this that Axiom XI and all the things which are asserted in plane geometry and trigonometry follow *absolutely* on F, L-lines taking the place of straight lines. Therefore the trigonometrical functions are to be accepted in the same sense as in Σ, and the circumference of the circle in F whose radius is the L-line $= r$, is $= 2\pi r$, and likewise $\odot r$ in F is $= \pi r^2$ (π being $\frac{1}{2}\odot 1$ in F, or 3.1415926...).



§22

If $\overline{AB}$ is the L of $\overline{AM}$ and C is in $\overline{AM}$, and the angle CAB formed from the straight line $A\overline{M}$ and the L-line $A\overline{B}$ is moved first along $A\overline{B}$ and then along $B\overline{A}$ to infinity, the path $\overline{CD}$ of C will be the L of $\overline{CM}$.

For (calling the latter l), let D be any point in $\overline{CD}$, $DN|||CM$, and B the point of L falling in $\overline{DN}$. $BN \doteq AM$, and $AC = BD$, and so also $DN \doteq CM$, and therefore D will be in l. But if D is in[2] l and $DN|||CM$, and B is the point of L common to it and

[1] [Proved in §9 only when one of the angles is a right angle.]

[2] [This is a new D. First he takes any point of the path of C and proves that it is on l, and then he takes any point on l and proves that it is a point of the path of C. Here $\overline{CD}$ is not necessarily straight.

$\overline{DN}$, then $AM \simeq BN$ and $CM \simeq DN$, whence it is clear that $BD = AC$, and D falls in the path of the point C and l and $\overline{CD}$ are the same. We designate such an l by $l \| L$.

§23

If the L-line $CDF \| ABE$ (§22), and $AB = BE$, and AM, $B\overline{N}$, and $E\overline{P}$ are axes, plainly $CD = DF$; and if any three points A, B, and E belong to AB and $AB = n \,.\, CD$, then will $AE = n \,.\, CF$, and therefore (plainly also for incommensurables AB, AE, and CD),

$$AB : CD = AE : CF,$$

and $AB : CD$ is *independent* of AB and *directly determined* by AC. Let this quantity be denoted by the capital letter (as X) of the same name as the small letter (as x) by which AC is denoted.

§24

Whatever be x and y, $Y = X^{\frac{y}{x}}$ (§23).

For one of the two letters x and y will be a multiple of the other (for example, y of x) or not.

If $y = nx$, let $x = AC = CG = GH$ etc., until is made $AH = y$. Then let $CD \| GK \| HL$. Then (§23)

$$X = AB : CD = CD : GK = GK : HL,$$

and so

$$\frac{AB}{HL} = \left(\frac{AB}{CD}\right)^n,$$

or

$$Y = X^n = X^{\frac{y}{x}}.$$

If x and y are multiples of i, say $x = mi$ and $y = ni$, then by the preceding $X = I^m$, $Y = I^n$, and therefore $Y = X^{\frac{n}{m}} = X^{\frac{y}{x}}$.

The same is easily extended to the case of incommensurability of x and y. But if $q = y - x$, clearly $Q = Y : X$.

Now it is manifest that in Σ for any x is $X = 1$; but in S, $X > 1$, and for *any* AB and ABE there is a $CDF \| ABE$ such that $CDF = AB$, whence $AMBN \equiv AMEP$[1] although the latter is a multiple of the former, which is indeed singular, but evidently does not prove the absurdity of S.

[1] [$AMBN$ means that portion of the plane that lies between the complete lines AM and BN, and $AMEP$ means that portion that lies between the complete lines AM and EP. See page 375 footnote 5.]

FERMAT

On Analytic Geometry

(Translated from the French by Professor Joseph Seidlin, Alfred College, Alfred, N. Y.)

The following extract is from Fermat's *Introduction aux Lieux Plans et Solides.* It appears in the the *Varia Opera Mathematica* of Fermat in 1679, and in the *Œuvres de Fermat*, ed. Tannery and Henry, Paris, 1896. It shows how clearly Fermat understood the connection between algebra and geometry. It will be observed that Fermat uses the terms "plane and solid loci" in an older sense, somewhat different from the one now recognized.

The French text will be found in the *Œuvres*, vol. III, pp. 85–96.

Introduction to Plane and Solid Loci

None can doubt that the ancients wrote on loci. We know this from Pappus, who, at the beginning of Book VII, affirms that Apollonius had written on plane loci and Aristæus on solid loci. But, if we do not deceive ourselves, the treatment of loci was not an easy matter for them. We can conclude this from the fact that, despite the great number of loci, they hardly formulated a single generalization, as will be seen later on. We therefore submit this theory to an apt and particular analysis which opens the general field for the study of loci.

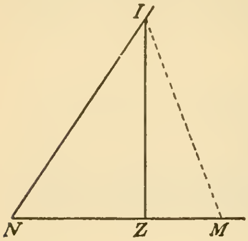
Whenever two unknown magnitudes appear in a final equation, we have a locus, the extremity of one of the unknown magnitudes describing a straight line or a curve. The straight line is simple and unique; the classes of curves are indefinitely many,—circle, parabola, hyperbola, ellipse, etc.

When the extremity of the unknown magnitude which traces the locus, follows a straight line or a circle, the locus is said to be plane; when the extremity describes a parabola, a hyperbola, or an ellipse, the locus is said to be solid....

It is desirable, in order to aid the concept of equation, to let the two unknown magnitudes form an angle, which usually we would suppose to be a right angle, with the position and the extreme point of one of the unknown magnitudes established. If neither of the two unknowns is greater than a quadratic, the locus will

be plane or solid, as can be clearly seen from the following:

Let NZM be a straight line of given position with point N fixed. Let NZ be the unknown quantity a and ZI (the line drawn to form the angle NZI) the other unknown quantity e.



If $da = be$, the point I will describe a line of fixed position. Indeed, we would have $\frac{b}{d} = \frac{a}{e}$. Consequently the ratio $a{:}e$ is given, as is also the angle at Z. Therefore both the triangle NIZ and the angle INZ are determined. But the point N and the position of the line NZ are given, and so the position of NI is determined. The synthesis is easy.

To this equation we can reduce all those whose terms are either known or combined with the unknowns a and e, which may enter simply or may be multiplied by given magnitudes.

$$z^{\text{II}} - da = be.$$

Suppose that $z^{\text{II}} = dr$. We then have

$$\frac{b}{d} = \frac{r - a}{e}.$$

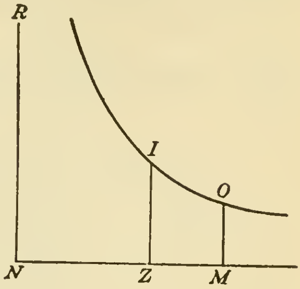
If we let $MN = r$, point M will be fixed and we shall have $MZ = r - a$.

The ratio $\frac{MZ}{ZI}$ therefore becomes fixed. With the angle at Z given, the triangle IZM will be determined, and in drawing MI it follows that this line is fixed. Thus point I will be on a line of determined position. A ike conclusion can be reached without difficulty for any equation containing the terms a or e.

Here is the first and simplest equation of a locus, from which all the loci of a straight line may be found; for example, the proposition 7 of Book I of Apollonius "On Plane Loci," which has since, however, found a more general expression and mode of construction. This equation yields the following interesting proposition: "Assume any number of lines of given position. From a given point draw lines forming given angles. If the sum of the products of the lines thus drawn by the given lines equals a given area, then the given point will trace a line of determined position."

We omit a great number of other propositions, which could be considered as corollaries to those of Apollonius.

The second species of equations of this kind are of the form $ae = z^{\text{II}}$, in which case point I traces a hyperbola. Draw NR parallel to ZI; through any point, such as M, on the line NZ, draw MO parallel to ZI. Construct the rectangle NMO equal in area to z^{II}. Through the point O, between the asymptotes NR, NM, describe a hyperbola; its position is determined and it will pass through point I, having assumed, as it were, ae,—that is to say the rectangle NZI,—equivalent to the rectangle NMO. To this equation we may reduce all those whose terms are in part constant, or in part contain a or e or ae.



If we let

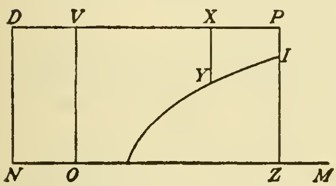
$$d^{\text{II}} + ae = ra + se$$

we obtain by fundamental principles $ra + se - ae = d^{\text{II}}$. Construct a rectangle of such dimensions as shall contain the terms $ra + se - ae$. The two sides will be $a - s$ and $r - e$, and their rectangle, $ra + se - ae - rs$.

If from d^{II} we subtract rs, the rectangle

$$(a - s)(r - e) = d^{\text{II}} - rs.$$

Take NO equal to s, and ND, parallel to ZI, equal to r. Through point D, draw DP parallel to NM; through point O, OV parallel to ND; prolong ZI to P.



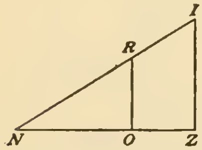
Since $NO = s$ and $NZ = a$, we have $a - s = OZ = VP$. Similarly, since $ND = ZP = r$ and $ZI = e$, we have $r - e = PI$. The rectangle $PV \times PI$ is therefore equal to the given area $d^{\text{II}} - rs$; the point I is therefore on a hyperbola having PV, VO as asymptotes.

If we take any point X, the parallel XY, and construct the rectangle $VXY = d^{\text{II}} - rs$, and through point Y we describe a hyperbola between the asymptotes PV, VO, it will pass through point I. The analysis and construction are easy in every case.

The following species of loci equations arises if we have $a^2 = e^2$ or if a^2 is in a given relation to e^2, or, again, if $a^2 + ae$ is in a given relation to e^2. Finally this type includes all the equations whose terms are of the second degree containing a^2, e^2, or ae. In all

these cases point I traces a straight line, which is easily demonstrated.

If the ratio $\frac{NZ^2 + NZ.ZI}{ZI^2}$ is given, and any parallel OR is drawn, then it is easy to show that $\frac{NO^2 + NO.OR}{OR^2}$ has the value

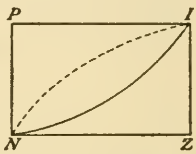


of the given ratio. The point I will therefore be on a line of determined position. The same will be true of all equations whose terms are either the squares of the unknowns or their product. It is needless to enumerate additional specific instances.

If to the squares of the unknowns, with or without their product, are added absolute terms or terms which are the products of one of the unknowns by a given magnitude, the construction is more difficult. We shall indicate the construction and give the proof for several cases.

If $a^2 = de$, point I is on a parabola.

Let NP be parallel to ZI; with NP as diameter, construct the parabola whose parameter is the given line d and whose ordinates are parallel to NZ. The point I will be on the parabola whose position is defined. In fact, it follows from the construction that the rectangle $d \times NP = PI^2$, that is, $d \times IZ = NZ^2$ and, consequent y, $de = a^2$.



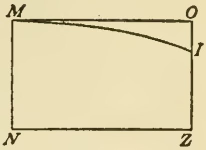
To this equation we can easily reduce all those in which, with a^2, appear the products of the given magnitudes and e, or with e^2 appear the products of the given magnitudes with a. The same would hold true were the equation to contain absolute terms.

If, however, $e^2 = da$, then, in the preceding figure, with N as vertex and with NZ as diameter, construct the parabola whose parameter is d and whose ordinates are parallel to the line NP. It is plain that the imposed condition is satisfied.

If we let $b^2 - a^2 = de$, we have $b^2 - de = a^2$. Divide b^2 by d; let $b^2 = dr$, and we have $dr - de = a^2$ or $d(r - e) = a^2$. We shall have reduced this equation to the former [—that is, $a^2 = de$,—] by replacing $r - e$ by e.

Let us assume MN (p. 393) parallel to ZI and equal to r; through the point M draw MO parallel to NZ. Point M and the position of the line MO are now given. It follows from the construction that $OI = r - e$. Therefore $d \times OI = NZ^2 = MO^2$.

The parabola drawn with M as vertex, diameter MN, d as parameter, and the ordinates parallel to NZ, satisfies the condition as is clearly shown by the construction.

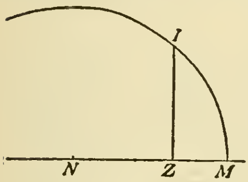


If $b^2 + a^2 = de$, we have $de - b^2 = a^2$, etc., as above. Similarly then we can construct all the equations containing a^2 and e.

But a^2 is often found with e^2 and with absolute terms. Let $b^2 - a^2 = e^2$.

The point I will be on a circle of determined position if the angle NZI is a right angle.

Assume MN equal to b. The circle described with N as center and with NM as radius will satisfy the condition. That is to say, that no matter which point I is taken, anywhere on the circumference, it is clear that ZI^2 (or e^2) will equal NM^2 (or b^2) $- NZ^2$ (or a^2).



To this equation may be reduced all those containing terms in a^2, e^2, and in a or e multiplied by given magnitudes, provided angle NZI be a right angle, and, moreover, that the coefficient of a^2 be equal to that of e^2.

Let

$$b^2 - 2da - a^2 = e^2 + 2re.$$

Adding r^2 to both sides and, thus replacing e by $e + r$, we have

$$r^2 + b^2 - 2da - a^2 = e^2 + r^2 + 2re.$$

Adding d^2 to $r^2 + b^2$, thus replacing a by $d + a$, and denoting the sum of the squares $r^2 + b^2 + d^2$ by p^2, we get

$$p^2 - d^2 - 2da - a^2 = r^2 + b^2 - 2da - a^2,$$

which leads to

$$p^2 - d^2 = r^2 + b^2.$$

If now we replace $a + d$ by a and $e + r$ by e, we shall have

$$p^2 - a^2 = e^2,$$

which equation is reduced to the preceding.

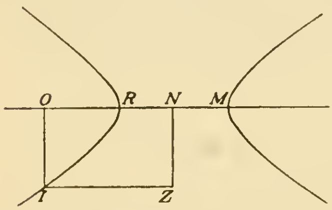
By like reasoning we are able to reduce all similar equations. Based on this method we have built up all of the propositions of the Second Book of Apollonius "On Plane Loci" and we have proved that the six first cases have loci for any points whatever, which is quite remarkable and which was probably unknown to Apollonius.

When $\frac{b^2 - a^2}{e^2}$ is a given ratio, the point I will be on an ellipse. Let MN equal b. With M as vertex, NM as diameter, and N as center describe an ellipse whose ordinates are parallel to ZI, so that the squares of the ordinates shall be in a given ratio to the product of the segment of the diameter. The point I will be on that ellipse. That is, $NM^2 - NZ^2$ is equal to the product of the segments of the diameter.

To this equation can be reduced all those in which a^2 is on one side of the equation and e^2 with an opposite sign and a different coefficient on the other side. If the coefficients are the same and the angle a right angle, the locus will be a circle, as we have said. If the coefficients are the same but the angle is not a right angle, the locus will be an ellipse.

Moreover, though the equations include terms which are products of a or e by given magnitudes, the reduction may nevertheless be made by the method which we have already employed.

If $(a^2 + b^2):e^2$ is a given ratio, the point I will be on a hyperbola. Draw NO parallel to ZI; let the given ratio be equal to $b^2:NR^2$. Point R will then be fixed. With R as vertex, RO as diameter, and N as center, construct an hyperbola whose ordinates are parallel to NZ, such that the product of the whole diameter (MR) by RO together with RO^2 shall be to OI^2 as $NR^2:b^2$. It follows, letting $MN = NR$, that $(MO \times OR + NR^2):(OI^2 + b^2)$ is equal to $NR^2:b^2$, the given ratio.



But

$$MO \times OR + NR^2 = NO^2 = ZI^2 = e^2$$

and

$$OI^2 + b^2 = NZ^2 \text{ (or } a^2\text{)} + b^2.$$

Therefore $e^2:(b^2 + a^2) = NR^2:b^2$ and, inverting, $(b^2 + a^2):e^2$ is the given ratio. Therefore point I is on an hyperbola of determined position.

By the scheme we have already employed we may reduce to this equation all those in which a^2 and e^2 are contained with given terms (separately) or with expressions involving the products of a or e by the given terms, and in which a^2 and e^2 have the same sign and appear on the opposite sides of the equation. If the signs were different the locus would be a circle or an ellipse.

The most difficult type of equation is that containing, along with a^2 and e^2, terms involving ae, other given magnitudes, etc.

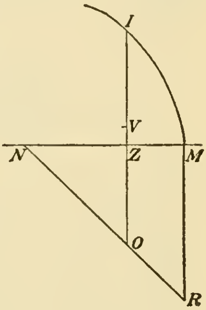
Let

$$b^2 - 2a^2 = 2ae + e^2.$$

Add a^2 to both sides so as to have $a + e$ as a factor of one of the members. Then

$$b^2 - a^2 = a^2 + 2ae + e^2.$$

Replace $a + e$ by, say, e; then, according to the preceding development, the circle MI will satisfy the equation; that is to say, MN^2 $(=b^2) - NZ^2$ $(=a^2) = ZI^2(=[a + e]^2)$. Letting $VI = NZ = a$, we have $ZV = e$.



In this problem, however, we are looking for the point V or the extremity of the line e. It is therefore necessary to find, and to indicate, the line upon which the point V is located.

Let MR be parallel to ZI and equal to MN. Draw NR which meets IZ, prolonged, at O. Since $MN = MR$, $NZ = ZO$. But $NZ = VI$; therefore, by addition, $VO = ZI$. Therefore $MN^2 - NZ^2 = VO^2$. But triangle NMR is known; therefore the ratio $NM^2{:}NR^2$ is given as are also the ratios $NZ^2{:}NO^2$ and $(MN^2 - NZ^2){:}(NR^2 - NO^2)$. But we have proved that $OV^2 = MN^2 - NZ^2$. Therefore the ratio $(NR^2 - NO^2){:}\,OV^2$ is known. But the points N and R are given, as well as the angle NOZ. Therefore, as we have just shown, point V is on an ellipse.

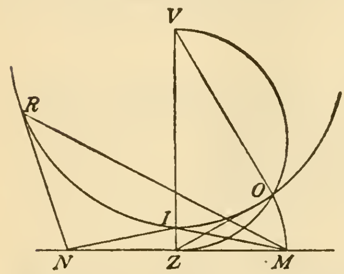
By analogous procedure we reduce to the preceding cases all the others in which along with the terms containing ae and a^2 or e^2 are also terms consisting of products of a and e by given magnitudes. The discussion of these different cases is very easy. The problem may always be solved by means of a triangle of known configuration.

We have therefore included in a brief and clear exposition all that the ancients have left unexplained concerning plane and solid loci. Consequently one can recognize at once which loci apply to all cases of the final proposition of Book I of Apollonius "On Plane Loci," and one can generally discover without great difficulty all which pertains to that matter.

As a culminating point to this treatise, we shall add a very interesting proposition of almost obvious simplicity:

"Given the position of any number of lines; if from some definite point lines be drawn forming given angles with the given lines, and the sum of the squares of all the segments is equal to a given area, the point will describe a solid locus of determined position."

A single example will suffice to indicate the general method of construction. Given two points *N* and *M*, required the locus of the points such that the sum of the squares of *IN*, *IM*, shall be in a given ratio to the triangle *INM*.



Let $NM = b$. Let e be the line ZI drawn at right angles to NM, and let a be the distance NZ. In accordance with fundamental principles, $(2a^2 + b^2 - 2ba + 2e^2) : be$ is a given ratio. Following in treatment the procedures previously explained we have the suggested construction.

Bisect *NM* at *Z*; erect at *Z* the perpendicular *ZV*; make the ratio $4ZV:NM$ equal to the given ratio. On *VZ* draw the semicircle *VOZ*, inscribe $ZO = ZM$, and draw *VO*. With *V* as center and *VO* as radius draw the circle *OIR*. If from any point *R* on this circle, we draw *RN*, *RM*, I say that $RN^2 + RM^2$ is in the given ratio to the triangle *RNM*.

The constructions of the theorems on loci could have been much more elegantly presented if this discovery had preceded our already old revision of the two books on plane loci. Yet, we do not regret this work, however precocious or insufficiently ripe it may be. In fact, there is for science a certain fascination in not exposing to posterity works which are as yet spiritually incomplete; the labor of the work at first simple and clumsy gains strength as well as stature through new inventions. It is quite important that the student should be able to discern clearly the progress which appears veiled as well as the spontaneous development of the science.

(Facing page 397.)

DESCARTES

On Analytic Geometry

(Translated from the French by Professor David Eugene Smith, Teachers College, Columbia University, New York City, and the late Marcia L. Latham, Hunter College, New York City.)[1]

René Descartes (1596–1650), philosopher, mathematician, physicist, soldier, and littérateur, published the first book that may properly be called a treatise on analytic geometry. This appeared as the third appendix to his *Discours de la méthode pour bien conduire sa raison et chercher la vérité dans les sciences*, which was published at Leyden in 1637. Pierre de Fermat (c. 1608–1665) had already conceived the idea as early as 1629, as is shown by a letter written by him to Roberval on Sept. 22, 1636, but he published nothing upon the subject. For the posthumous publication see p. 389.

The following extract constitutes the first eight pages of the first edition (pages 297–304, inclusive) of the *Discours*.

La Geometrie

Book I

Problems the Construction of Which Requires Only Straight Lines and Circles

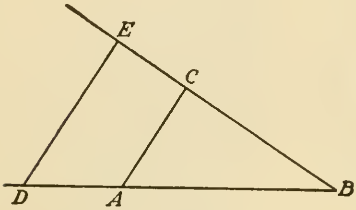
Any problem in geometry can easily be reduced to such terms that a knowledge of the lengths of certain straight lines is sufficient for its construction.[2] Just as arithmetic consists of only four or five operations, namely, addition, subtraction, multiplication, division, and the extraction of roots, which may be considered a kind of division, so in geometry, to find required lines it is merely necessary to add or subtract other lines; or else, taking one line which I shall call unity in order to relate it as closely as possible

[1] From the edition of *La Géométrie* published in facsimile and translation by The Open Court Publishing Co., Chicago, 1925, and here reprinted with the permission of the publisher.

[2] [Large collections of problems of this nature are contained in the following works: Vincenzo Riccati and Girolamo Saladino, *Institutiones Analyticae*, Bologna, 1765; Maria Gaetana Agnesi, *Istituzioni Analitiche*, Milan, 1748; Claude Rabuel, *Commentaires sur la Géométrie de M. Descartes*, Lyons, 1730; and other books of the same period or earlier.]

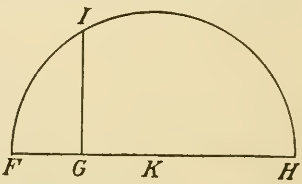
to numbers,[1] and which can in general be chosen arbitrarily, and having given two other lines, to find a fourth line which shall be to one of the given lines as the other is to unity (which is the same as multiplication); or, again, to find a fourth line which is to one of the given lines as unity is to the other (which is equivalent to division); or, finally, to find one, two, or several mean proportionals between unity and some other line (which is the same as extracting the square root, cube root, etc., of the given line). And I shall not hesitate to introduce these arithmetical terms into geometry, for the sake of greater clearness.

For example, let AB be taken as unity, and let it be required to multiply BD by BC. I have only to join the points A and C, and draw DE parallel to CA; then BE is the product of BD and BC.



If it be required to divide BE by BD, I join E and D, and draw AC parallel to DE; then BC is the result of the division.

If the square root of GH is desired, I add, along the same straight line, FG equal to unity; then, bisecting FH at K, I describe the circle FIH about K as a center, and draw from G a perpendicular and extend it to I, and GL is the required root. I do not speak here of cube roots, or other roots, since I shall speak more conveniently of them later.



Often it is not necessary thus to draw the lines on paper, but it is sufficient to designate each by a single letter. Thus, to add the lines BD and GH, I call one a and the other b, and write $a + b$. Then $a - b$ will indicate that b is subtracted from a; ab that a is multiplied by b; $\frac{a}{b}$ that a is divided by b; aa or a^2 that a is multiplied by itself; a^3 that this result is multiplied by a, and so on, indefinitely.[2] Again, if I wish to extract the square root of $a^2 + b^2$,

[1] [Van Schooten, in his Latin edition of 1683, has this note: "*Per unitatem intellige lineam quandam determinatam qua ad quamvis reliquarum linearum talem relationem babeat, qualem unitas ad certum aliquem numerum.*"]

[2] [Descartes uses a^3, a^4, a^5, a^6, and so on, to represent the respective powers of a, but he uses both aa and a^2 without distinction. For example, he often has $aabb$, but he also uses $3a^2/4b^2$.]

I write $\sqrt{a^2 + b^2}$; if I with to extract the cube root of $a^3 - b^3 + ab^2$, I write $\sqrt[3]{a^3 - b^3 + ab^2}$, and similarly for other roots.[1] Here it must be observed that by a^2, b^3, and similar expressions, I ordinarily mean only simple lines, which, however, I name squares, cubes, etc., so that I may make use of the terms employed in algebra.

It should also be noted that all parts of a single line should always be expressed by the same number of dimensions, provided unity is not determined by the conditions of the problem. Thus, a^3 contains as many dimensions as ab^2 or b^3, these being the component parts of the line which I have called $\sqrt[3]{a^3 - b^3 + ab^2}$. It is not, however, the same thing when unity is determined, because unity can always be understood, even when there are too many or too few dimensions; thus, if it be required to extract the cube root of $a^2b^2 - b$, we must consider the quantity a^2b^2 divided once by unity, and the quantity b multiplied twice by unity.[2]

Finally, so that we may be sure to remember the names of these lines, a separate list should always be made as often as names are assigned or changed. For example, we may write, $AB = 1$, that is AB equal to 1;[3] $GH = a$, $BD = b$, and so on.

If, then, we wish to solve any problem, we first suppose the solution already effected, and give names to all the lines that seem needful for its construction,—to those that are unknown as well as to those that are known. Then, making no distinction between known and unknown lines, we must unravel the difficulty in any way that shows most naturally the relations between these lines, until we find it possible to express a single quantity in two ways.[4] This will constitute an equation, since the terms of one of these two expressions are together equal to the terms of the other.

---

[1] [Descartes writes: $\sqrt{C.a^3 - b^3 + abb}$.]

[2] [Descartes seems to say that each term must be of the third degree, and that therefore we must conceive of both a^2b^2 and b as reduced to the proper dimension.]

[3] [Van Schooten adds "seu unitati," p. 3. Descartes writes, $AB \propto 1$. He seems to have been the first to use this symbol. Among the few writers who followed him, was Hudde (1633–1704). It is very commonly supposed that $\propto$ is a ligature representing the first two letters (or diphthong) of "aequale." See, for example, M. Aubry's note in W. R. Ball's *Recréations Mathématiques et Problèmes des Temps Anciens et Modernes*, French edition, Paris, 1909, Part III, p. 164. See also F. Cajori, *Hist. of Math. Notations*, vol. I, p. 301.]

[4] [That is, we must solve the remaining simultaneous equations.]

We must find as many such equations as there are supposed to be unknown lines;[1] but if, after considering everything involved, so many cannot be found, it is evident that the question is not entirely determined. In such a case we may choose arbitrarily lines of known length for each unbroken line to which there corresponds no equation.

If there are several equations, we must use each in order, either considering it alone or comparing it with the others, so as to obtain a value for each of the unknown lines; and so we must combine them until there remains a single unknown line which is equal to some known line, or whose square, cube, fourth power, fifth power, sixth power, etc., is equal to the sum or difference of two or more quantities, one of which is known, while the others consist of mean proportionals between unity and this square, or cube, or fourth power, etc., multiplied by other known lines. I may express this as follows:

$$z = b,$$

or

$$z^2 = -az + b^2,$$

or

$$z^3 = az^2 + b^2z - c^3,$$

or

$$z^4 = az^3 - c^3z + d^4, \text{ etc.}$$

That is, z, which I take for the unknown quantity, is equal to b; or, the square of z is equal to the square of b diminished by a multiplied by z; or, the cube of z is equal to a multiplied by the square of z, plus the square of b multiplied by z, diminished by the cube of c; and similarly for the others.

Thus, all the unknown quantities can be expressed in terms of a single quantity, whenever the problem can be constructed by means of circles and straight lines, or by conic sections, or even by some other curve of degree not greater than the third or fourth.

But I shall not stop to explain this in more detail, because I should deprive you of the pleasure of mastering it yourself, as well

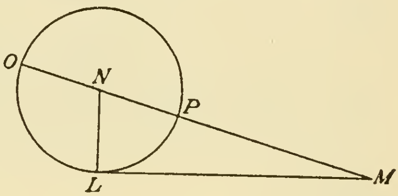
---

[1] [Van Schooten (p. 149) gives two problems to illustrate this statement. Of these, the first is as follows: Given a line segment AB containing any point C, required to produce AB to D so that the rectangle $AD.DB$ shall be equal to the square on CD. He lets $AC = a$, $CB = b$, and $BD = x$. Then $AD = a + b + x$, and $CD = b + x$, whence $ax + bx + x^2 = b^2 + 2bx + x^2$ and $x = \dfrac{b^2}{a - b}$.]

as of the advantage of training your mind by working over it, which is in my opinion the principal benefit to be derived from this science. Because, I find nothing here so difficult that it cannot be worked out by any one at all familiar with ordinary geometry and with algebra, who will consider carefully all that is set forth in this treatise.

I shall therefore content myself with the statement that if the student, in solving these equations, does not fail to make use of division wherever possible, he will surely reach the simplest terms to which the problem can be reduced.

And if it can be solved by ordinary geometry, that is, by the use of straight lines and circles traced on a plane surface, when the last equation shall have been entirely solved there will remain at most only the square of an unknown quantity, equal to the product of its root by some known quantity, increased or diminished by some other quantity also known.[1] Then this root or unknown line can easily be found. For example, if I have $z^2 = az + b^2$,[2] I construct a right triangle NLM with one side LM, equal to b, the square root of the known quantity b^2, and the other side, LN, equal to $\frac{1}{2}a$, that is to half the other known quantity which was multiplied by z, which I suppose to be the unknown line. Then prolonging MN, the hypotenuse[3] of this triangle, to O, so that NO is equal to NL, the whole line OM is the required line z. This is expressed in the following way:[4]



$$z = \tfrac{1}{2}a + \sqrt{\tfrac{1}{4}a^2 + b^2}.$$

But if I have $y^2 = -ay + b^2$, where y is the quantity whose value is desired, I construct the same right triangle NLM, and

[1] [That is, an expression of the form $z^2 = az \pm b$. "Esgal a ce qui se produit de l'Addition, ou soustraction de sa racine multiplée par quelque quantité connue, & de quelque autre quantité aussy connue."]

[2] [Descartes proposes to show how a quadratic may be solved geometrically.]

[3] [Descartes says "prolongeant MN la baze de ce triangle," because the hypotenuse was commonly taken as the base in earlier times.]

[4] [From the figure $OM.PM = LM^2$. If $OM = z$, $PM = z - a$, and since $LM = b$, we have $z(z - a) = b^2$ or $z^2 = az + b^2$. Again, $MN = \sqrt{\frac{1}{4}a^2 + b^2}$, whence $OM = z = ON + MN = \frac{1}{2}a + \sqrt{\frac{1}{4}a^2 + b^2}$. Descartes ignores the second root, which is negative.]

on the hypotenuse MN lay off NP equal to NL, and the remainder PM is y, the desired root. Thus I have

$$y = -\tfrac{1}{2}a + \sqrt{\tfrac{1}{4}a^2 + b^2}$$

In the same way if I had

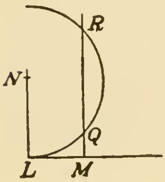
$$x^4 = -ax^2 + b^2,$$

PM would be x^2 and I should have

$$x = \sqrt{-\tfrac{1}{2}a + \sqrt{\tfrac{1}{4}a^2 + b^2}},$$

and so for other cases.

Finally, if I have $z^2 = az - b^2$, I make NL equal to $\tfrac{1}{2}a$ and LM equal to b as before; then, instead of joining the points M and N, I draw MQR parallel to LN, and with N as a center describe a circle through L cutting MQR in the points Q and R; then z, the line sought, is either MQ or MR, for in this case it can be expressed in two ways, namely,



$$z = \tfrac{1}{2}a + \sqrt{\tfrac{1}{4}a^2 - b^2},$$

and

$$z = \tfrac{1}{2}a - \sqrt{\tfrac{1}{4}a^2 - b^2}.$$

And if the circle described about N and passing through L neither cuts nor touches the line MQR, the equation has no root, so that we may say that the construction of the problem is impossible.

These same roots can be found by many other methods. I have given these very simple ones to show that it is possible to construct all the problems of ordinary geometry by doing no more than the little covered in the four figures that I have explained.[1] This is one thing which I believe the ancient mathematicians did not observe, for otherwise they would not have put so much labor into writing so many books in which the very sequence of the propositions shows that they did not have a sure method of finding all,[2] but rather gathered those propositions on which they had happened by accident.

[1] [It will be seen that Descartes considers only three types of the quadratic equation in z, namely $z^2 + az - b^2 = 0$, $z^2 - az - b^2 = 0$, and $z^2 - az + b^2 = 0$. It thus appears that he has not been able to free himself from the old traditions to the extent of generalizing the meaning of the coefficients,—as negative and fractional as well as positive. He does not consider the type $z^2 + az + b^2 = 0$, because it has no positive roots.]

[2] ["Qu'ils n'ont point eu la vraye methode pour les trouuer toutes."]

POHLKE'S THEOREM

(Translated from the German by Professor Arnold Emch, University of Illinois, Urbana, Ill.)

Karl Pohlke was born in Berlin on January 28, 1810, and died there November 27, 1876. He taught in various engineering schools, closing his work in the Technische Hochschule in Charlottenburg. In his *Darstellende Geometrie* (Berlin, 1859; 4th ed., 1876, p. 109) is found "the principal theorem of axonometry," now generally known as Pohlke's Theorem. It is here translated as an important piece of source material, but without proof, the demonstration being available in an article by Arnold Emch, *American Journal of Mathematics*, vol. 40 (1918). On the general development of orthogonal axonometry see F. J. Obenrauch, *Geschichte der darstellenden und projectiven Geometrie*, Brünn, 1897, pp. 385 seq.

Three segments of arbitrary length a_1x_1, a_1y_1, a_1z_1, which are drawn in a plane from a point a_1 under arbitrary angles, form a parallel projection of three equal segments ax, ay, az from the origin on three perpendicular coordinate axes; however, only one of the segments $a_1x_1, \ldots$, or one of the angles may vanish.

RIEMANN

On Riemann's Surfaces and Analysis Situs

(Translated from the German by Dr. James Singer, Princeton University, Princeton, N. J.)

Georg Friedrich Bernhard Riemann (1826–1866) was born at Breselenz in Hannover and died at Selasca on his third trip to Italy. He studied theology at Göttingen and also attended some mathematical lectures there. He soon gave up theology for mathematics and studied under Gauss and Stern. In 1847 he went to Berlin, drawn by the fame of Dirichlet, Jacobi, Steiner, and Eisenstein. He returned to Göttingen in 1850 to study physics under Weber and there he received his doctorate the following year. He became a Privatdozent at Göttingen in 1854, a professor in 1857, and in 1859 succeeded Dirichlet as ordinary professor. His work on the differential equations of physics, a series of lectures edited by Hattendorf and later by H. Weber, is still a standard textbook. In prime numbers also he opened a new field. The first section of this article is a translation of part of Riemann's paper entitled "Allgemeine Voraussetzungen und Hülfsmittel für die Untersuchung von Functionen unbeschränkt veränderlicher Grössen," which appeared in Crelle's *Journal für reine und angewandte Mathematik*, Bd. 54, 1857, pp. 103–104. The second division is a translation of part of his paper, "Lehrsätze aus der Analysis Situs für die Theorie der Integrale von zweigliedrigen vollständigen Differentialien," which appeared in the same issue, pp. 105–110. The papers can also be found in his *Mathematische Werke* collected by H. Weber, first edition, pp. 83–89; second edition, pp. 90–96.

The importance of these two contributions of Riemann can scarcely be exaggerated. Thanks to the Riemann surface the theory of single-valued analytic functions of one variable can largely be extended to multiple-valued functions. Riemann introduced his surface for the purpose of studying algebraic functions, in which field it plays a fundamental part. In Riemann's work we find also the real beginning of modern Analysis Situs, which in a larger sense was however created by Poincaré (1895).

1. For many investigations, especially in the investigation of algebraic and Abelian functions, it is advantageous to represent geometrically the modes of branching of a multivalued function in the following way: We imagine another surface spread over the (x, y)-plane and coincident with it (or an infinitely thin body spread over the plane) which extends as far and only as far as the function is defined. By a continuation of this function this surface will

be likewise further extended. In a portion of the plane in which there exist two or more continuations of the function, the surface will be double or multifold; it will consist there of two or more sheets, each one of which represents one branch of the function. Around a branch point of the function a sheet of the surface will be continued into another, so that in the neighborhood of such a point the surface can be considered as a helicoid with its axis perpendicular to the (x, y)-plane at this point, and with infinitely small pitch. If the function takes on again its original value after several revolutions of z around a branch point (for example, as $(z - a)^{\frac{m}{n}}$, where m and n are relatively prime numbers, after n revolutions of z around a), we must then of course assume that the topmost sheet of the surface is continued into the lowermost by means of the remaining ones. The multiple-valued function has only one definite value for each point of the surface representing its modes of branching and therefore can be regarded as a fully determined function of the position in this surface.

. . . .

2. In the investigation of functions which arise from the integration of total differentials several theorems belonging to Analysis Situs are almost indispensable. This name, used by Leibniz, although perhaps not entirely with the same significance, may well designate a part of the theory of continuous entities which treats them not as existing independently of their positions and measurable by one another but, on the contrary, entirely disregarding the metrical relations, investigates their local and regional properties. While I propose to present a treatment entirely free from metric considerations, I will here present in a geometric form only the theorems necessary for the integration of two-termed total differentials.

. . . .

If upon a surface F two systems of curves, a and b, together completely bound a part of this surface, then every other system of curves which together with a, completely bounds a part of F also constitutes with b the complete boundary of a part of the surface; which part is composed of both of the first partitions of the surface joined along a (by addition or subtraction, according as they lie upon opposite or upon the same side of a). Both systems of curves serve equally well for the complete boundary

of a part of F and can be interchanged as far as the satisfying of this requirement.[1]

If upon the surface F there can be drawn n closed curves $a_1, a_2, \ldots, a_n$ *which neither by themselves nor with one another completely bound a part of this surface F, but with whose aid every other closed curve does form the complete boundary of a part of F, the surface is said to be* $(n + 1)$*-fold connected.*

This character of the surface is independent of the choice of the system of curves $a_1, a_2, \ldots, a_n$ since any other n closed curves $b_1, b_2, \ldots, b_n$ which are not sufficient to bound completely a part of this surface, do likewise completely bound a part of F when taken together with any other closed curve.

Indeed, since b_1 completely bounds a part of F when taken together with curves a, one of these curves a can be replaced by b_2 and the remaining curves a. Therefore, any other curve, and consequently also b_2, together with b_1 and these $n - 1$ curves a is sufficient for the complete boundary of a part of F, and hence one of these $n - 1$ curves a can be replaced by b_1, b_2 and the remaining $n - 2$ curves a. If, as assumed, the curves b are not sufficient for the complete boundary of a part of F, this process can clearly be continued until all the a's have been replaced by the b's.

---

[1] [Note by H. Weber.] The theorem stated here needs to be somewhat restricted and made more precise, as was pointed out by Tonelli (*Atti della R. accademia dei Lincei*, Ser. II, vol. 2, 1875. In an extract from the *Nachrichten der Gesselschaft der Wissenschaften zu Göttingen*, 1875.)

If the system of curves a completely bounds a part of the surface F when taken together with a system of curves b as well as with a second system of curves c, it is generally necessary, in order that the systems of curves b and c taken together likewise bound a part of the surface, that no subset of the curves a together with b or with c already bounds a part of the surface. The part of the surface bounded by the systems of curves b, c which, even when the parts of the surface a, b, and a, c are simple, can consist of several separate pieces, are described by Tonelli in the following fashion: It consists of the totality of the parts of the surface a, b, and a, c when those parts which are bounded by the curves a are taken away from the parts common to both of these surface partitions.

The example given by Tonelli of a closed five-fold connected double anchor ring bounded by a point illustrates this relation and makes it intuitive.

These remarks have no influence on the use which Riemann makes of this theorem for the definition of the $(n + 1)$-fold connectivity, since the system here denoted by a always consists of only one curve, namely the curve a which is replaced by b.

By means of a crosscut,—i. e., a line lying in the interior of the surface and going from a boundary point to a boundary point,—an $(n + 1)$-fold connected surface F can be changed into an n-fold connected one, F'. The parts of the boundary arising from the cutting play the role of boundary even during the further cutting so that a crosscut can pass through no point more than once but can end in one of its earlier points.

Since the lines $a_1, a_2, \ldots, a_n$ are not sufficient for the complete boundary of a part of F, if one imagines F cut up by these lines, then the piece of the surface lying on the right of a_n as well as that lying on the left must contain boundary elements other than the lines a and which belong therefore to the boundary of F. One can therefore draw a line in the one as well as the other of these pieces of surface not cutting the curves a from a point of a_n to the boundary of F. Both of these two lines q' and q'' taken together then constitute a crosscut q of the surface F which satisfies the requirement.

Indeed, on the surface F' arising from this crosscut of F the curves $a_1, a_2, \ldots, a_{n-1}$ are closed curves lying in the interior of F' which are not sufficient to bound a part of F and hence also not a part of F'. However, every other closed curve l lying in the interior of F' constitutes with them the complete boundary of a part of F'. For the line l forms with a complex of the lines $a_1, a_2, \ldots, a_n$ the complete boundary of a part f of F. However, it can be shown that a_n cannot occur in the boundary of f; because then, according as f lies on the left or right side of a_n, q' or q'' would go from the interior of f to a boundary point of F, hence to a point lying outside of f, and therefore would cut the boundary of f contrary to the hypothesis that l as well as the lines a, excepting for the point of intersection of a_n and q, always lie in the interior of F'.

The surface F' into which F is decomposed by the crosscut q is therefore n-fold connected, as required.

It shall now be shown that the surface F is changed into an n-fold connected one F' by any crosscut p which does not decompose it into separate pieces. If the pieces of surface adjacent to the two sides of the crosscut p are connected, a line b can be drawn from one side of p through the interior of F' back to the starting point on the other side. This line b forms a line in the interior of F leading back into itself; and since the crosscut issuing from it on both sides goes to a boundary point, b cannot constitute the com-

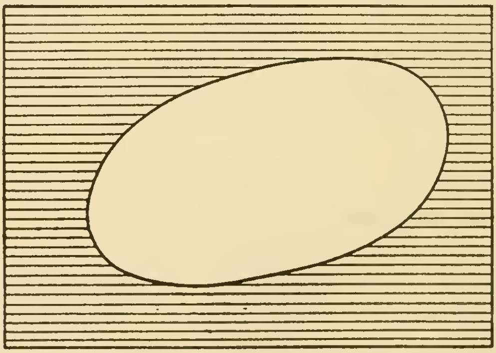
plete boundary of either of the two pieces of surface into which it separates F. We can therfore replace one of the curves a by the curve b and each of the remaining $n - 1$ curves a by a curve in the interior of F' and the curve b, if necessary; wherefrom we can deduce by the same means as above the proof that F' is n-fold connected.

An $(n + 1)$*-fold connected surface will therefore be changed into an* n*-fold connected one by means of any crosscut which does not separate it into pieces.*

The surface arising from a crosscut can be divided again by a new crosscut, and after n repetitions of this operation an $(n + 1)$-fold connected surface will be changed into a simply connected one by means of n successive non-interesting crosscuts. To apply these considerations to a surface without boundary, a closed surface, we must change it into a bounded one by the specialization of an arbitrary point; so that the first division is made by means of this point and a crosscut beginning and ending in it, hence by a closed curve. For example, the surface of an anchor ring, which is 3-fold connected, will be changed into a simply connected surface by means of a closed curve and a crosscut.

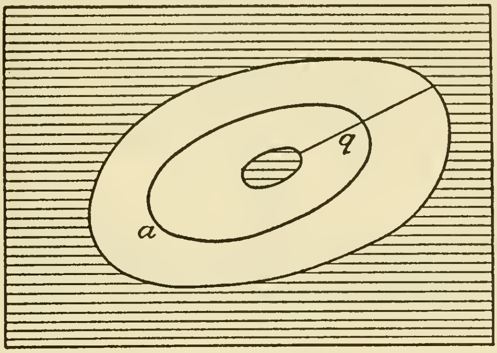
Simply-connected Surface

It will be decomposed into parts by any crosscut, and any closed curve in it constitutes the complete boundary of a part of the surface.



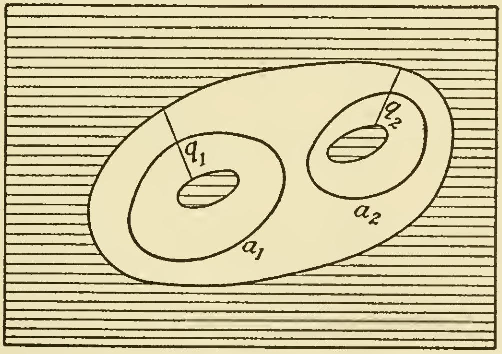
Doubly-connected Surface

It will be reduced to a simply-connected one by any crosscut q that does not disconnect it. Any closed curve in it can, with the aid of a, constitute the complete boundary of a part of the surface.

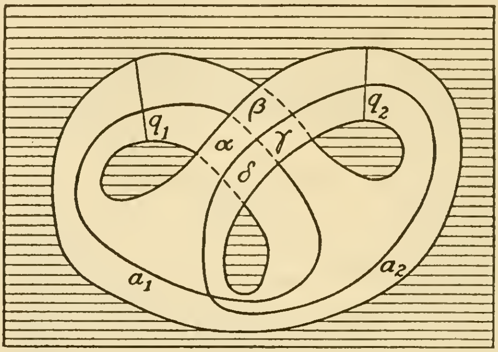


Triply-connected Surface

In this surface any closed curve can constitute the complete boundary of a part of the surface with the aid of the curves a_1 and a_2. It is decomposed into a doubly connected surface by any crosscut that does not disconnect it and into a simply connected one by two such crosscuts, q_1 and q_2.



This surface is double in the regions α, β, γ, δ of the plane. The arm of the surface containing a_1 is imagined as lying under the other and is therefore represented by dotted lines.



RIEMANN

On the Hypotheses which Lie at the Foundations of Geometry

(Translated from the German by Professor Henry S. White, Vassar College, Poughkeepsie, N. Y.)

For a biographical sketch of Riemann see page 404.

The paper here translated is Riemann's *Probe-Vorlesung*, or formal initial lecture on becoming Privat-Docent. It is extraordinary in scope and originality and it paved the way for the now current theories of hyperspace and relativity. It was read on the 10th of June, 1854, for the purpose of Riemann's "Habilitation" with the philosophical faculty of Göttingen. This explains the form of presentation, in which analytic investigations could be only indicated; some elaborations of them are to be found in the "*Commentatio mathematica, qua respondere tentatur quaestioni ab Illma Academia Parisiensi propositae*" etc., and in the appendix to that paper. It appears in vol. XIII of the *Abhandlungen* of the Royal Society of Sciences of Göttingen.

Plan of the Investigation

It is well known that geometry presupposes not only the concept of space but also the first fundamental notions for constructions in space as given in advance. It gives only nominal definitions for them, while the essential means of determining them appear in the form of axioms. The relation of these presuppositions is left in the dark; one sees neither whether and in how far their connection is necessary, nor a priori whether it is possible.

From Euclid to Legendre, to name the most renowned of modern writers on geometry, this darkness has been lifted neither by the mathematicians nor by the philosophers who have labored upon it. The reason of this lay perhaps in the fact that the general concept of multiply extended magnitudes, in which spatial magnitudes are comprehended, has not been elaborated at all. Accordingly I have proposed to myself at first the problem of constructing the concept of a multiply extended magnitude out of general notions of quantity. From this it will result that a multiply extended magnitude is susceptible of various metric relations and that space accordingly constitutes only a particular case of a

triply extended magnitude. A necessary sequel of this is that the propositions of geometry are not derivable from general concepts of quantity, but that those properties by which space is distinguished from other conceivable triply extended magnitudes can be gathered only from experience. There arises from this the problem of searching out the simplest facts by which the metric relations of space can be determined, a problem which in nature of things is not quite definite; for several systems of simple facts can be stated which would suffice for determining the metric relations of space; the most important for present purposes is that laid down for foundations by Euclid. These facts are, like all facts, not necessary but of a merely empirical certainty; they are hypotheses; one may therefore inquire into their probability, which is truly very great within the bounds of observation, and thereafter decide concerning the admissibility of protracting them outside the limits of observation, not only toward the immeasurably large, but also toward the immeasurably small.

I. *The Concept of n-fold Extended Manifold*

While I now attempt in the first place to solve the first of these problems, the development of the concept of manifolds multiply extended, I think myself the more entitled to ask considerate judgment inasmuch as I have had little practise in such matters of a philosophical nature, where the difficulty lies more in the concepts than in the construction, and because I have not been able to make use of any preliminary studies whatever aside from some very brief hints which Privy Councillor Gauss has given on the subject in his second essay on biquadratic residues and in his Jubilee booklet, and some philosophical investigations of Herbart.

1

Notions of quantity are possible only where there exists already a general concept which allows various modes of determination. According as there is or is not found among these modes of determination a continuous transition from one to another, they form a continuous or a discrete manifold; the individual modes are called in the first case points, in the latter case elements of the manifold. Concepts whose modes of determination form a discrete manifold are so numerous, that for things arbitrarily given there can always be found a concept, at least in the more highly developed languages, under which they are comprehended

(and mathematicians have been able therefore in the doctrine of discrete quantities to set out without scruple from the postulate that given things are to be considered as all of one kind); on the other hand there are in common life only such infrequent occasions to form concepts whose modes of determination form a continuous manifold, that the positions of objects of sense, and the colors, are probably the only simple notions whose modes of determination form a multiply extended manifold. More frequent occasion for the birth and development of these notions is first found in higher mathematics.

Determinate parts of a manifold, distinguished by a mark or by a boundary, are called quanta. Their comparison as to quantity comes in discrete magnitudes by counting, in continuous magnitude by measurement. Measuring consists in superposition of the magnitudes to be compared; for measurement there is requisite some means of carrying forward one magnitude as a measure for the other. In detault of this, one can compare two magnitudes only when the one is a part of the other, and even then one can only decide upon the question of more and less, not upon the question of how many. The investigations which can be set on foot about them in this case form a general part of the doctrine of quantity independent of metric determinations, where magnitudes are thought of not as existing independent of position and not as expressible by a unit, but only as regions in a manifold. Such inquiries have become a necessity for several parts of mathematics, namely for the treatment of many-valued analytic functions, and the lack of them is likely a principal reason why the celebrated theorem of Abel and the contributions of Langrange, Pfaff, and Jacobi to the theory of differential equations have remained so long unfruitful. For the present purpose it will be sufficient to bring forward conspicuously two points out of this general part of the doctrine of extended magnitudes, wherein nothing further is assumed than what was already contained in the concept of it. The first of these will make plain how the notion of a multiply extended manifold came to exist; the second, the reference of the determination of place in a given manifold to determinations of quantity and the essential mark of an n-fold extension.

2

In a concept whose various modes of determination form a continuous manifold, if one passes in a definite way from one mode

of determination to another, the modes of determination which are traversed constitute a simply extended manifold and its essential mark is this, that in it a continuous progress is possible from any point only in two directions, forward or backward. If now one forms the thought of this manifold again passing over into another entirely different, here again in a definite way, that is, in such a way that every point goes over into a definite point of the other, then will all the modes of determination thus obtained form a doubly extended manifold. In similar procedure one obtains a triply extended manifold when one represents to oneself that a double extension passes over in a definite way into one entirely different, and it is easy to see how one can prolong this construction indefinitely. If one considers his object of thought as variable instead of regarding the concept as determinable, then this construction can be characterized as a synthesis of a variability of $n + 1$ dimensions out of a variability of n dimensions and a variability of one dimension.

3

I shall now show how one can conversely split up a variability, whose domain is given, into a variability of one dimension and a variability of fewer dimensions. To this end let one think of a variable portion of a manifold of one dimension,—reckoning from a fixed starting-point or origin, so that its values are comparable one with another—which has for every point of the given manifold a definite value changing continuously with that point; or in other words, let one assume within the given manifold a continuous function of place, and indeed a function such that it is not constant along any portion of this manifold. Every system of points in which the function has a constant value constitutes now a continuous manifold of fewer dimensions than that which was given. By change in the value of the function these manifolds pass over, one into another, continuously; hence one may assume that from one of them all the rest emanate, and this will come about, speaking generally, in such a way that every point of one passes over into a definite point of the other. Exceptional cases, and it is important to investigate them,—can be left out of consideration here. By this means the fixing of position in the given manifold is referred to the determination of one quantity and the fixing of position in a manifold of fewer dimensions. It is easy now to show that this latter has $n - 1$ dimensions if the given manifold

was n-fold extended. Hence by repetition of this procedure, to n times, the fixing of position in an n-dimensional manifold is reduced to n determinations of quantities, and therefore the fixing of position in a given manifold is reduced, whenever this is possible, to the determination of a finite number of quantities. There are however manifolds in which the fixing of position requires not a finite number but either an infinite series or a continuous manifold of determinations of quantity. Such manifolds are constituted for example by the possible determinations of a function for a given domain, the possible shapes of a figure in space, et cetera.

II. Relations of Measure, of Which an n-dimensional Manifold is Susceptible, on the Assumption that Lines Possess a Length Independent of Their Position; that is, that Every Line Can Be Measured by Every Other

Now that the concept of an n-fold extended manifold has been constructed and its essential mark has been found to be this, that the determination of position therein can be referred to n determinations of magnitude, there follows as second of the problems proposed above, an investigation into the relations of measure that such a manifold is susceptible of, also into the conditions which suffice for determining these metric relations. These relations of measure can be investigated only in abstract notions of magnitude and can be exhibited connectedly only in formulae; upon certain assumptions, however, one is able to resolve them into relations which are separately capable of being represented geometrically, and by this means it becomes possible to express geometrically the results of the calculation. Therefore if one is to reach solid ground, an abstract investigation in formulae is indeed unavoidable, but its results will allow an exhibition in the clothing of geometry. For both parts the foundations are contained in the celebrated treatise of Privy Councillor Gauss upon curved surfaces.

1

Determinations of measure require magnitude to be independent of location, a state of things which can occur in more than one way. The assumption that first offers itself, which I intend here to follow out, is perhaps this, that the length of lines be independent of their situation, that therefore every line be measurable by every

other. If the fixing of the location is referred to determinations of magnitudes, that is, if the location of a point in the n-dimensional manifold be expressed by n variable quantities x_1, x_2, x_3, and so on to x_n, then the determination of a line will reduce to this, that the quantities x be given as functions of a single variable. The problem is then, to set up a mathematical expression for the length of lines, and for this purpose the quantities x must be thought of as expressible in units. This problem I shall treat only under certain restrictions, and limit myself first to such lines as have the ratios of the quantities dx—the corresponding changes in the quantities x—changing continuously; one can in that case think of the lines as laid off into elements within which the ratios of the quantities dx may be regarded as constant, and the problem reduces then to this: to set up for every point a general expression for a line-element which begins there, an expression which will therefore contain the quantities x and the quantities dx. In the second place I now assume that the length of the line-element, neglecting quantities of the second order, remains unchanged when all its points undergo infinitely small changes of position; in this it is implied that if all the quantities dx increase in the same ratio, the line-element likewise changes in this ratio. Upon these assumptions it will be possible for the line-element to be an arbitrary homogeneous function of the first degree in the quantities dx which remains unchanged when all the dx change sign, and in which the arbitrary constants are continuous functions of the quantities x. To find the simplest cases, I look first for an expression for the $(n - 1)$-fold extended manifolds which are everywhere equally distant from the initial point of the line-element, that is, I look for a continuous function of place, which renders them distinct from one another. This will have to diminish or increase from the initial point out in all directions; I shall assume that it increases in all directions and therefore has a minimum in that point. If then its first and second differential quotients are finite, the differential of the first order must vanish and that of the second order must never be negative; I assume that it is always positive. This differential expression of the second order accordingly remains constant if ds remains constant, and increases in squared ratio when the quantities dx and hence also ds all change in the same ratio. That expression is therefore = const. ds^2, and consequently ds = the square root of an everywhere positive entire homogeneous function of the second degree in

quantities dx having as coefficients continuous functions of the quantities x. For space this is, when one expresses the position of a point by rectangular coordinates, $ds = \sqrt{\Sigma(dx)^2}$; space is therefore comprised under this simplest case. The next case in order of simplicity would probably contain the manifolds in which the line-element can be expressed by the fourth root of a differential expression of the fourth degree. Investigation of this more general class indeed would require no essentially different principles, but would consume considerable time and throw relatively little new light upon the theory of space, particularly since the results cannot be expressed geometrically. I limit myself therefore to those manifolds in which the line-element is expressed by the square root of a differential expression of the second degree. Such an expression one can transform into another similar one by substituting for the n independent variables functions of n new independent variables. By this means however one cannot transform every expression into every other; for the expression contains $n \cdot \frac{n+1}{2}$ coefficients which are arbitrary functions of the independent variables; but by introducing new variables one can satisfy only n relations (conditions), and so can make only n of the coefficients equal to given quantities. There remain then $n \cdot \frac{n-1}{2}$ others completely determined by the nature of the manifold that is to be represented, and therefore for determining its metric relations $n \cdot \frac{n-1}{2}$ functions of position are requisite. The manifolds in which, as in the plane and in space, the line-element can be reduced to the form $\sqrt{\Sigma(dx)^2}$ constitute therefore only a particular case of the manifolds under consideration here. They deserve a particular name, and I will therefore term *flat* these manifolds in which the square of the line-element can be reduced to the sum of squares of total differentials. Now in order to obtain a conspectus of the essential differences of the manifolds representable in this prescribed form it is necessary to remove those that spring from the mode of representation, and this is accomplished by choosing the variable quantities according to a definite principle.

2

For this purpose suppose the system of shortest lines emanating from an arbitrary point to have been constructed. The position

of an undetermined point will then be determinable by specifying the direction of that shortest line in which it lies and its distance, in that line, from the starting-point; and it can therefore be expressed by the ratios of the quantities dx^0, that is the limiting ratios of the dx at the starting point of this shortest line and by the length s of this line. Introduce now instead of the dx^0 such linear expressions $d\alpha$ formed from them, that the initial value of the square of the line-element equals the sum of the squares of these expressions, so that the independent variables are: the quantity s and the ratios of quantities $d\alpha$. Finally, set in place of the $d\alpha$ such quantities proportional to them, $x_1, x_2, \ldots, x_n$, that the sum of their squares $= s^2$. After introducing these quantities, the square of the line-element for indefinitely small values of x becomes $= \Sigma(dx)^2$, and the term of next order in that $(ds)^2$ will be equal to a homogeneous expression of the second degree in the $n\frac{n-1}{2}$ quantities $(x_1dx_2 - x_2dx_1)$, $(x_1dx_3 - x_3dx_1), \ldots$, that is, an indefinitely small quantity of dimension four; so that one obtains a finite magnitude when one divides it by the square of the indefinitely small triangle-area in whose vertices the values of the variables are $(0, 0, 0, \ldots)$, $(x_1, x_2, x_3, \ldots)$, $(dx_1, dx_2, dx_3, \ldots)$. This quantity retains the same value, so long as the quantities x and dx are contained in the same binary linear forms, or so long as the two shortest lines from the values 0 to the values x and from the values 0 to the values dx stay in the same element of surface, and it depends therefore only upon the place and the direction of that element. Plainly it is $= 0$ if the manifold represented is flat, that is if the square of the line-element is reducible to $\Sigma(dx)^2$, and it can accordingly be regarded as the measure of the divergence of the manifold from flatness in this point and in this direction of surface. Multiplied by $-\frac{3}{4}$ it becomes equal to the quantity which Privy Councillor Gauss has named the measure of curvature of a surface.

For determining the metric relations of an n-fold extended manifold representable in the prescribed form, in the foregoing discussion $n \cdot \frac{n-1}{2}$ functions of position were found needful; hence when the measure of curvature in every point in $n \cdot \frac{n-1}{2}$ surface-directions is given, from them can be determined the metric relations of the manifold, provided no identical relations exist

among these values, and indeed in general this does not occur. The metric relations of these manifolds that have the line-element represented by the square root of a differential expression of the second degree can thus be expressed in a manner entirely independent of the choice of the variable quantities. A quite similar path to this goal can be laid out also in case of the manifolds in which the line-element is given in a less simple expression; *e. g.*, as the fourth root of a differential expression of the fourth degree. In that case the line-element, speaking generally, would no longer be reducible to the form of a square root of a sum of squares of differential expressions; and therefore in the expression for the square of the line-element the divergence from flatness would be an indefinitely small quantity of the dimension two, while in the former manifolds it was indefinitely small of the dimension four. This peculiarity of the latter manifolds may therefore well be called flatness in smallest parts. The most important peculiarity of these manifolds, for present purposes, on whose account solely they have been investigated here, is however this, that the relations of those doubly extended can be represented geometrically by surfaces, and those of more dimensions can be referred to those of the surfaces contained in them; and this requires still a brief elucidation.

3

In the conception of surfaces, along with the interior metric relations, in which only the length of the paths lying in them comes into consideration, there is always mixed also their situation with respect to points lying outside them. One can abstract however from external relations by carrying out such changes in the surfaces as leave unchanged the length of lines in them; *i. e.*, by thinking of them as bent in any arbitrary fashion,—without stretching—and by regarding all surfaces arising in this way one out of another as equivalent. For example, arbitrary cylindrical or conical surfaces are counted as equivalent to a plane, because they can be formed out of it by mere bending, while interior metric relations remain unchanged; and all theorems regarding them—the whole of planimetry—retain their validity; on the other hand they count as essentially distinct from the sphere, which cannot be converted into a plane without stretching. According to the above investigation in every point the interior metric relations of a doubly extended manifold are characterized by the measure

of curvature if the line-element can be expressed by the square root of a differential expression of the second degree, as is the case with surfaces. An intuitional significance can be given to this quantity in the case of surfaces, namely that it is the product of the two curvatures of the surface in this point; or also, that its product into an indefinitely small triangle-area formed of shortest lines is equal to half the excess of its angle-sum above two right angles, when measured in radians. The first definition would presuppose the theorem that the product of the two radii of curvature is not changed by merely bending a surface; the second, the theorem that at one and the same point the excess of the angle-sum of an indefinitely small triangle above two right angles is proportional to its area. To give a tangible meaning to measure of curvature of an n-dimensional manifold at a given point and in a surface direction passing through that point, it is necessary to start out from the principle that a shortest line, originating in a point, is fully determined when its initial direction is given. According to this, a determinate surface is obtained when one prolongs into shortest lines all the initial directions going out from a point and lying in the given surface element; and this surface has in the given point a determinate measure of curvature, which is also the measure of curvature of the n-dimensional manifold in the given point and the given direction of surface.

4

Now before applications to space some considerations are needful regarding flat manifolds in general, *i. e.*, regarding those in which the square of the line-element is representable by a sum of squares of total differentials.

In a flat n-dimensional manifold the measure of curvature at every point is in every direction zero; but by the preceding investigation it suffices for determining the metric relations to know that at every point, in $n \cdot \frac{n-1}{2}$ surface directions whose measures of curvature are independent of one another, that measure is zero. Manifolds whose measure of curvature is everywhere zero may be regarded as a particular case of those manifolds whose curvature is everywhere constant. The common character of those manifolds of constant curvature can also be expressed thus: that the figures lying in them can be moved without stretching. For it is evident that the figures in them could not be pushed along and

rotated at pleasure unless in every point the measure of curvature were the same in all directions. Upon the other hand, the metric relations of the manifold are completely determined by the measure of curvature. About any point, therefore, the metric relations in all directions are exactly the same as about any other point, and so the same constructions can be carried out from it, and consequently in manifolds with constant curvature every arbitrary position can be given to the figures. The metric relations of these manifolds depend only upon the value of the measure of curvature, and it may be mentioned, with reference to analytical presentation, that if one denotes this value by α, the expression for the line element can be given the form

$$\frac{1}{1 + \frac{\alpha}{4}\sqrt{\Sigma dx^2}}$$

5

Consideration of surfaces with constant measure of curvature can help toward a geometric exposition. It is easy to see that those surfaces whose curvature is positive will always permit themselves to be fitted upon a sphere whose radius is unity divided by the square root of the measure of curvature; but to visualize the complete manifold of these surfaces one should give to one of them the form of a sphere and to the rest the form of surfaces of rotation which touch it along the equator. Such surfaces as have greater curvature than this sphere will then touch the sphere from the inner side and take on a form like that exterior part of the surface of a ring which is turned away from the axis (remote from the axis); they could be shaped upon zones of spheres having a smaller radius, but would reach more than once around. Surfaces with lesser positive measure of curvature will be obtained by cutting out of spherical surfaces of greater radius a portion bounded by two halves of great circles, and making its edges adhere together. The surface with zero curvature will be simply a cylindrical surface standing upon the equator; the surfaces with negative curvature will be tangent to this cylinder externally and will be formed like the inner part of the surface of a ring, the part turned toward the axis.

If one thinks of these surfaces as loci for fragments of surface movable in them, as space is for bodies, then the fragments are movable in all these surfaces without stretching. Surfaces with

positive curvature can always be formed in such wise that those fragments can be moved about without even bending, namely as spherical surfaces, not so however those with negative curvature. Beside this independence of position shown by fragments of surface, it is found in the surface with zero curvature that direction is independent of position, as is not true in the rest of the surfaces.

III. Application to Space

1

Following these investigations concerning the mode of fixing metric relations in an n-fold extended magnitude, the conditions can now be stated which are sufficient and necessary for determining metric relations in space, when it is assumed in advance that lines are independent of position and that the linear element is representable by the square root of a differential expression of the second degree; that is if flatness in smallest parts is assumed.

These conditions in the first place can be expressed thus: that the measure of the curvature in every point is equal to zero in three directions of surface; and therefore the metric relations of the space are determined when the sum of the angles in a triangle is everywhere equal to two right angles.

In the second place if one assumes at the start, like Euclid, an existence independent of situation not only for lines but also for bodies, then it follows that the measure of curvature is everywhere constant; and then the sum of the angles in all triangles is determined as soon as it is fixed for one triangle.

In the third place, finally, instead of assuming the length of lines to be independent of place and direction, one might even assume their length and direction to be independent of place. Upon this understanding the changes in place or differences in position are complex quantities expressible in three independent units.

2

In the course of preceding discussions, in the first place relations of extension (or of domain) were distinguished from those of measurement, and it was found that different relations of measure were conceivable along with identical relations of extension. Then were sought systems of simple determinations of measure by means of which the metric relations of space are completely deter-

mined and of which all theorems about such relations are a necessary consequence. It remains now to examine the question how, in what degree and to what extent these assumptions are guaranteed by experience. In this connection there subsists an essential difference between mere relations of extension and those of measurement: in the former, where the possible cases form a discrete manifold the declarations of experience are indeed never quite sure, but they are not lacking in exactness; while in the latter, where possible cases form a continuum, every determination based on experience remains always inexact, be the probability that it is nearly correct ever so great. This antithesis becomes important when these empirical determinations are extended beyond the limits of observation into the immeasurably great and the immeasurably small; for the second kind of relations obviously might become ever more inexact, beyond the bounds of observation, but not so the first kind.

When constructions in space are extended into the immeasurably great, unlimitedness must be distinguished from infiniteness; the one belongs to relations of extension, the other to those of measure. That space is an unlimited, triply extended manifold is an assumption applied in every conception of the external world; by it at every moment the domain of real perceptions is supplemented and the possible locations of an object that is sought for are constructed, and in these applications the assumption is continually being verified. The unlimitedness of space has therefore a greater certainty, empirically, than any experience of the external. From this, however, follows in no wise its infiniteness, but on the contrary space would necessarily be finite, if one assumes that bodies are independent of situation and so ascribes to space a constant measure of curvature, provided this measure of curvature had any positive value however small. If one were to prolong the elements of direction, that lie in any element of surface, into shortest lines (geodetics), one would obtain an unlimited surface with constant positive measure of curvature, consequently a surface which would take on, in a triply extended manifold, the form of a spherical surface, and would therefore be finite.

3

Questions concerning the immeasurably large area, for the explanation of Nature, useless questions. Quite otherwise is it however with questions concerning the immeasurably small.

Knowledge of the causal connection of phenomena is based essentially upon the precision with which we follow them down into the infinitely small. The progress of recent centuries in knowledge of the mechanism of Nature has come about almost solely by the exactness of the syntheses rendered possible by the invention of Analysis of the infinite and by the simple fundamental concepts devised by Archimedes, Galileo, and Newton, and effectively employed by modern Physics. In the natural sciences however, where simple fundamental concepts are still lacking for such syntheses, one pursues phenomen into the spatially small, in order to perceive causal connections, just as far as the microscope permits. Questions concerning spatial relations of measure in the indefinitely small are therefore not useless.

If one premise that bodies exist independently of position, then the measure of curvature is everywhere constant; then from astronomical measurments it follows that it cannot differ from zero; at any rate its reciprocal value would have to be a surface in comparison with which the region accessible to our telescopes would vanish. If however bodies have no such non-dependence upon position, then one cannot conclude to relations of measure in the indefinitely small from those in the large. In that case the curvature can have at every point arbitrary values in three directions, provided only the total curvature of every metric portion of space be not appreciably different from zero. Even greater complications may arise in case the line element is not representable, as has been premised, by the square root of a differential expression of the second degree. Now however the empirical notions on which spatial measurements are based appear to lose their validity when applied to the indefinitely small, namely the concept of a fixed body and that of a light-ray; accordingly it is entirely conceivable that in the indefinitely small the spatial relations of size are not in accord with the postulates of geometry, and one would indeed be forced to this assumption as soon as it would permit a simpler explanation of the phenomena.

The question of the validity of the postulates of geometry in the indefinitely small is involved in the question concerning the ultimate basis of relations of size in space. In connection with this question, which may well be assigned to the philosophy of space, the above remark is applicable, namely that while in a discrete manifold the principle of metric relations is implicit in the notion of this manifold, it must come from somewhere else

in the case of a continuous manifold. Either then the actual things forming the groundwork of a space must constitute a discrete manifold, or else the basis of metric relations must be sought for outside that actuality, in colligating forces that operate upon it.

A decision upon these questions can be found only by starting from the structure of phenomena that has been approved in experience hitherto, for which Newton laid the foundation, and by modifying this structure gradually under the compulsion of facts which it cannot explain. Such investigations as start out, like this present one, from general notions, can promote only the purpose that this task shall not be hindered by too restricted conceptions, and that progress in perceiving the connection of things shall not be obstructed by the prejudices of tradition.

This path leads out into the domain of another science, into the realm of physics, into which the nature of this present occasion forbids us to penetrate.

"Riemann, who was logically the immediate predecessor of Einstein, brought in a new idea of which the importance was not perceived for half a century. He considered that geometry ought to start from the infinitesimal, and depend upon integration for statements about finite lengths, areas, or volumes. This requires, inter alia, the replacement of the straight line by the geodesic: the latter has a definition depending upon infinitesimal distances, while the former has not. The traditional view was that, while the length of a curve could, in general, only be defined by integration, the length of the straight line between two points could be defined as a whole, not as the limit or a sum of little bits. Riemann's view was that a straight line does not differ from a curve in this respect. Moreover, measurement, being performed by means of bodies, is a physical operation, and its results depend for their interpretation upon the laws of physics. This point of view has turned out to be of very great importance. Its scope has been extended by the theory of relativity, but in essence it is to be found in Riemann's dissertation." (Bertrand Russell, *The Analysis of Matter*, p. 21, New York, 1927, Harcourt, Brace and Company, Quoted by permission of the publishers.

MONGE

On the Purpose of Descriptive Geometry

(Translated from the French by Professor Arnold Emch, University of Illinois, Urbana, Ill.)

Gaspard Monge (1746–1818) was the son of an itinerant tradesman. At twenty-two he was professor of mathematics in the military school at Mézières and finally held a similar position in the École Polytechnique in Paris. He is known chiefly for his elaboration of descriptive geometry, a theory which had been suggested by Frézier in 1738. He lectured upon the subject at Paris in "l'an 3 de la République" (1794–1795) and his *Géométrie Descriptive* was published in "l'an 7" (1798–1799). He had already laid the foundations for the theory when teaching at Mézières, and on January 11, 1775 he had presented a memoir before the Académie des Sciences in which he made use of two planes of projection. He was one of the leaders in the foundation and organization of the École Normale and the École Polytechnique. The following brief quotation from his treatise (5th ed., Paris, 1927, pp. 1–2) will serve to set forth the purpose which he had in view and which the government guarded as a secret for some years because of its value in the construction of fortifications:

Descriptive geometry has two objects: the first is to establish methods to represent on drawing paper which has only two dimensions,—namely, length and width,—all solids of nature which have three dimensions,—length, width, and depth,—provided, however, that these solids are capable of rigorous definition.

The second object is to furnish means to recognize accordingly an exact description of the forms of solids and to derive thereby all truths which result from their forms and their respective positions.

REGIOMONTANUS

On the Law of Sines for Spherical Triangles

(Translated from the Latin by Professor Eva M. Sanford, College for Women, Western Reserve University, Cleveland, Ohio.)

Johann Müller (1436–1476), known as Regiomontanus, was the first to write a treatise devoted wholly to trigonometry. This appeared in manuscript about 1464, and had the title *De triangulis omnimodis.* The completeness of this work may be judged from the author's treatment of the Law of Sines for Spherical Triangles, a theorem which was probably of his own invention.

In every right-angled triangle, the ratio of the sine of each side to the sine of the angle which it subtends is the same.[1]

Given the triangle *abg* having the angle *b* a right angle. I say that the ratio of the sine of the side *ab* to the sine of the angle *agb* is the same as the ratio of the sine of the side *bg* to the sine of the angle *bag*, and also as the ratio of the sine of the side *ag* to the sine of the angle *abg*, which we shall prove as follows.

It is inevitable that each of the angles *a* and *g* is a right angle, or that one or other of them is a right angle, or that neither is a right angle. If each of them is a right angle, then, by hypothesis, the point *a* is the pole of the circle *bg*, and moreover the point *b* is the pole of the circle *ag* and *g* is the pole of the circle *ab*. Thus, by definition, each of the three arcs will measure[2] its respective angle. Therefore, the sine of any one of the three sides will be the same as that of the angle opposite, and accordingly the sine of each side has the same ratio to the sine of its respective angle, this ratio being that of equality.

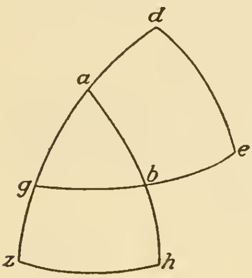
If, however, but one of the angles *a* and *g* is a right angle, let this one be the angle *g*. Since the hypothesis made *b* a right angle also, then, on this supposition, *a* is the pole of the circle *bg*, and each of the arcs *ba* and *ag* is a fourth of a great circle. Thus by definition, each of the arcs *ab*, *bg*, and *ga* will determine the size

[1] [*De triangulis omnimodis*, Lib. IIII, XVI, pp. 103–105, Nürnberg, 1533. The proof as given in this work has been divided into paragraphs for greater clarity in the translation.]

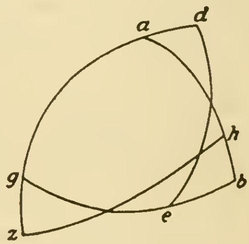
[2] [Literally "will determine the size of."]

of its respective angle, and the sine of any side will be the same as that of the corresponding angle by applying the definition of the sine of the angle. It will then be evident that the sine of each side has the same ratio, namely that of equality, to the sine of its corresponding angle.

But if neither of the angles *a* and *g* be a right angle, no one of the three sides will be a quadrant of a great circle, but they will be found in three-fold variety.[1] If each of the angles *a* and *g* should be acute, each of the arcs *ab* and *bg* will be less than a quadrant, and accordingly the arc *ag* will be less than a fourth of a great circle. Then let the arc *ga* be produced[2] toward *a* until it becomes the quadrant *gd*, and taking the chord which is the side of a great square[3] as a radius and the point *g* as a center, describe a great circle cutting the arc *gb* produced in the point *e*. Finally, let the arc *ag* be extended to the point *z* thus obtaining the quadrant *az* whose chord, swung about the pole *a* generates a circle which meets the arc *ab* extended in the point *h*. We have drawn a diagram illustrating these conditions.



But if each of the angles be obtuse, each of the arcs *ab* and *gb* will exceed a quadrant, and we know that the arc *ag* is less than a quadrant. Therefore, prolonging the arc *ag* on both sides as before until the fourth arc *gd* is formed and the arc *az* also, let two great circles be described with the centers at *g* and *a*. The circumference of the one described with *g* as a center will necessarily cut the arc *gb*, which is greater than a quadrant. Let this happen at the point *e*. The other circle described with *a* as a center will cut the arc *ab* at the point *h*. Thus another figure will be produced.

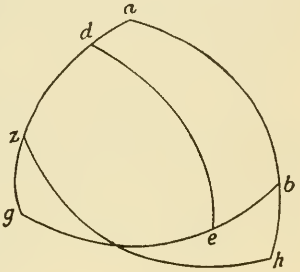


[1] [That is, all three sides will be less than a quadrant, or *ab* and *bg* will each be greater and *ag* less, or *bg* and *ag* will each be greater than a quadrant and *ab* will be less. It should be noted that Regiomontanus uses the letters in his diagrams in the order in which they appear in the Greek alphabet which is a natural outcome of his familiarity with mathematical classics in Greek.]

[2] [Literally "increased."]

[3] [Literally, "costa quadrati magni." This is evidently the chord of the quadrant of a great circle, the pole being used as a center in describing the circle on the sphere.]

But if one of the angles a and g is obtuse and the other acute, let a be obtuse and let the other be acute. Then, according to the cases cited, each of the arcs bg and ga is greater than a quadrant, but the arc ab is less than a quadrant. Therefore let two quadrants gd and az, which share the arc dz, be cut off from the arc ag. Then the circumference of a circle described as before with g as a pole will cut the arc bg which is greater than a quadrant. Let e be this point of intersection. Moreover the circumference of the circle described about a will not cut the arc ab, since this arc is less than a quadrant, but it will meet it if it is prolonged sufficiently, as at h. Therefore when neither of the angles a and g is a right angle, although we use a triple diagram, yet a single syllogism will result.



Since the two circles gd and ge meet obliquely,[1] and since two points are marked on the circumference of the circle gd with the perpendiculars ab and de drawn at these points, then according to the preceding demonstration the ratio of the sine of the arc ga to the sine of the arc ab will be as the sine of the arc gd to the sine of the arc de, and, by interchanging these terms,[2] the ratio of the sine ga to the sine gd will be that of the sine ab to the sine de. In like fashion, the two circles az and ah meet obliquely, and two points g and z are marked on the circumference of the circle az from which are drawn two perpendicular arcs gb and zh. Therefore, according to the foregoing proofs, the ratio of the sine ag to the sine gb is as that of the sine az to the sine zh; and by alternation, the sine ag is to the sine az as the sine gb is to the sine zh. Moreover, the sine ag is to the sine az as the sine ga is to the sine gd. Each of the arcs az and ga is a quadrant. Therefore the sine of the side ab has the same ratio to the sine de as the sine of the side gb has to the sine zh, which is that of the sine of the side ag to the sine of the quadrant. Moreover, the sine de is the sine of the angle agb, for the arc de measures the angle agb with the point g acting as the pole of the circle de. In like manner, the sine zh is the sine of the angle bag. Furthermore, the sine of the quadrant is the sine of a right angle, therefore the ratio of the sine of the side ab to the sine of the angle agb, and that of the sine of the side

[1] [Literally, "are inclined toward each other."]

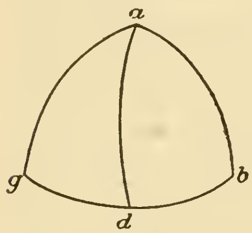
[2] [Literally, "by permuting the terms."]

bg to the sine of the angle *bag*, and also the ratio of the sine of the side *ag* to the sine of the right angle *abg* are one and the same, which was to be shown.

In every triangle, not right-angled, the sines of the sides have the same ratio as the sines of the angles opposite.[1]

The statement which the preceding proposition demonstrated for right-angled triangles may be proved for triangles that are not right-angled. Suppose that the triangle *abg* has no right angle. I say that the ratio of the sine of the side *ab* to the sine of the angle *g*, and that of the sine of the side *bg* to the sine of the angle *a*, and of the sine of the side *ga* to the sine of the angle *b* are one and the same.

I draw a perpendicular *ad* from *a* cutting the arc *bg* if it remains inside the triangle, or meeting the arc *bg* opportunely prolonged if it falls outside the triangle but being coterminus with neither *ab* nor *ag*; for in such a case, one of the angles *b* and *g* would be considered to be a right angle which our hypothesis has stated is not a right angle. Therefore; let it fall first within the triangle, marking out two triangles *abd* and *agd*. According to the preceding proof, but alternating the terms, the ratio of the sine *ab* to the sine *ad* is the same as that of the sine of the angle *adb*, a right angle, to that of the angle *abd*. But by the same previous proof, the ratio of the sine *ad* to the sine *ag* is the same as that of the sine of the angle *agd* to the sine of a right angle *adg*, since the sine of the angle *adg* is the same as that of the angle *adb* and since each of them is a right angle. Then[2] the sine of *ab* will be to the sine of *ag* as the sine of the angle *agb* is to the sine of the angle *abg*; and by alternation, the sine of the side *ab* will be to the sine of the angle *agb* as the sine of the side *ag* is to the sine of the angle *abg*.



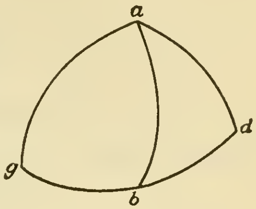
Finally, you will conclude that the ratio of the sine of the side *bg* to the sine of the angle *bag* is the same, if from one of the vertices *b* or *g* you draw an arc perpendicular to the side opposite it.

But if the perpendicular *ad* falls outside the triangle, thus changing the figure a little, let us seek the original syllogism; for reasoning by alternation from the preceding proof, the sine *ab* will be to the sine *ad* as the sine of the right angle *adb* is to the sine

[1] [Lib. IIII, XVII.]

[2] ["By reason of the equal indirect proportion."]

of the angle *abd*. Likewise, the sine of *ad* will be to the sine *ag* as the sine of the angle *agb* is to the sine of the right angle *adg*. Therefore, the sine of the side *ab* will be to the sine of the side *ag* as the sine of the angle *agb* is to the sine of the angle *abd*. Moreover, the sine of the angle *abd* is also the sine of the angle *abg* by common knowledge.[1] Therefore the sine *ab* is to the sine *ag* as the sine of the angle *agb* is to the sine of the angle *abd*, and thus, changing the terms, the sine of the side *ab* is to the sine of the angle *agb* as the sine of the side *ag* is to the sine of the angle *abg*. Finally, we shall prove that this is the ratio of the sine of the side *bg* to the sine of the angle *bag*, by the method which we have used above. Therefore the statement which was demonstrated in these theorems in regard to right-angled and non-right-angled triangles, respectively, we are at last free to state in general in regard to all triangles of whatever sort they may be, and we shall now consider step by step the great and jocund fruits which this study is to yield.



[1] ["Per communem scientiam" a direct translation from the Greek name for axiom.]

REGIOMONTANUS

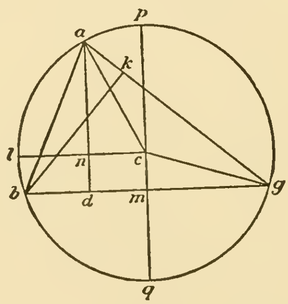
On the Relations of the Parts of a Triangle

(Translated from the Latin by Professor Vera Sanford, Western Reserve University, Cleveland, Ohio.)

Regiomontanus is the Latin name assumed by Johann Müller (1436–1476), being derived from his birthplace, Königsberg, in Lower Franconia. In a block-book almanac prepared by him his name appears as Magister Johann van Kunsperck. He was known in Italy, where he spent some years, as Joannes de Monteregio. He wrote *De triangulis omnimodis* c. 1464, but it was not printed until 1533. It was the first work that may be said to have been devoted solely to trigonometry. The following extract is from this work, lib. II, p. 58. In it Regiomontanus shows the relations of the parts of a triangle, and from it is easily derived the formula which, in our present symbols, would appears as $\Delta = \frac{1}{2}bc \sin A$.

XXVI

Given the area of a triangle and the rectangle[1] of the two sides, then the angle opposite the base will either be known or with the known angle will equal two right angles.[2]



Using again the diagrams of the preceding proposition, if the perpendicular *bk* meeting the line *ag* falls outside the triangle, then by the first case, the ratio of *bk* to *ba* will be known, and so

[1] [*I.e.*, the product.]

[2] [This enables one to find sin *A*, having given the area of the triangle and the product *bc*. Regiomontanus, however, does not seem to have changed this into the form: Given *A*, *b*, and *c* to find the area. The theorem determines the acute angle at the vertex, whether this be interior or exterior to the triangle.]

by the angle of this first figure, we shall assume *bak* as known, accordingly the angle *bag* with the known angle *bak* will equal two right angles. But if the perpendicular *bk* falls inside the triangle as is seen in the third diagram [the one here shown] of the preceeding proposition, then as before *ab* will have a known ratio to *bk*, and therefore the angle *bak* or *bag* will be known. But if the perpendicular *bk* coincides with the side *ab*, the angle *bag* must have been a right angle and therefore must be known, which indeed happens when the area of the proposed triangle equals that of the rectangle which is inclosed by the two sides.

PITISCUS

On the Laws of Sines and Cosines

(Translated from the Latin by Professor Jekuthiel Ginsburg, Yeshiva College, New York City.)

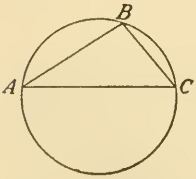
Bartholemäus Pitiscus (1561–1613), a German clergyman, wrote the first satisfactory textbook on trigonometry, and the first book to bear this title, —the *Trigonometriae sive de dimensione triangulorum libri quinque* (Frankfort, 1595, with later editions in 1599, 1600, 1608, and 1612, and an English edition in 1630). The selections here translated are from the 1612 edition, pages 95 and 105, and set forth the laws of sines and cosines. The translation makes use of modern symbols.

Fragment I

The ratio of the sides of a triangle to each other is the same as the ratio of the sines of the opposite angles.

The sines are halves of the [corresponding] chords. The sides of a triangle have the same ratio as the chords of the opposite angles, hence the ratio of the sides will be equal to the ratio of the sines, because the ratio of the whole quantity to another whole quantity is the same as the ratio of a half to a half, according to proposition 19 of Book 2, and it lies in the nature of the thing itself.[1]

The sides of the plane triangle will be the chords of the opposite angles or of the arcs by half of which the angles are measured.



Thus: If the circle ABC be circumscribed around the triangle ABC, the side AB will be the chord of angle ACB; that is, of the arc AB which measures the angle ACB. The side BC will be the chord of the angle BAC; that is, the chord of the arc BC which measures the angle BAC. Similarly the side AC will be the chord of the angle ABC; that is, of the arc AC which determines the angle ABC.

[1] [To prove this Pitiscus uses the circumscribed circle in the following way.]

Hence the side AB has the same ratio to the side BC as the chord of the angle ACB to the chord of the angle BAC which was to be proved.[1]

. . . .

Fragment II

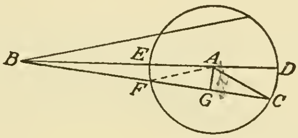
When the three sides of an oblique triangle are given, the segments made by the altitude drawn from the vertex of the greatest angle are given.[2]. . . . .

Subtract the square of one of the lateral sides of a triangle from the sum of the squares of the other two. Divide the remainder by twice the base and you will get the segment between the altitude and the other lateral side.[3]

---

[1] [According to Tropfke in his *Geschichte der Elementar-Mathematik* (V, p. 74) there were two methods of proving the Law of Sines: one used by Vieta (1540–1603) and traced back to Levi Ben Gerson (1288–1344), who was the first to formulate it in the West; the other to Nasîr ed-din al-Tûsî (1201–1274) and used by Regiomontanus, Pitiscus, and others. This is the method here given. It is equivalent to the modern method of expressing the sides a, b, c, as $2r \sin A$, $2r \sin B$, $2r \sin C$ respectively.]

[2] [In the ΔABC, AG is $\perp$ to BC.



To find CG, Pitiscus describes a circle with a radius AC and uses the known proportion

$$BC:BD = BE:BF,$$

in which $BC = a$, $AC = b$, $BA = c$. Then $BD = BA + AD = BA + AC = c + b$. Also, $BE = BA - AE = c - AC = c - b$, and $BF = BC - CF = a - 2x$.

Hence

$$a:c + b = c - b:a - 2x,$$

or

$$c^2 - b^2 = a^2 - 2ax;$$

therefore

$$2ax = a^2 + b^2 - c^2$$

$$x = \frac{a^2 + b^2 - c^2}{2a}$$

From this he derives the scholium which follows.]

[3] [From this there is only one step to the general form of the Law of Cosines. Pitiscus did not make that step, perhaps because he considered it self-evident; but he used the theorem above given in exactly the same way as we now use the Law of Cosines; that is, he used it in finding the values of the angles from the given sides.]

PITISCUS

On Bürgi's Method of Trisecting an Arc

(Translated from the Latin by Professor Jekuthiel Ginsburg, Yeshiva College, New York City.)

Jobst Bürgi's (1552–1632) solution of the equation used in the trisection of an arc was given by Bartholomeus Pitiscus (1561–1613) in his *Trigonometria*, (1595; 1612 edition, pages 50–54). Whether Bürgi obtained it from Arabic sources or discovered it independently is an interesting question that has not as yet been answered satisfactorily.

The material in the translated "fragments" is interesting on account of the bearing it has on questions of both algebra and trigonometry. The explanation consists of two fragments, one of which is introductory to the other.

Fragment 1 [p. 38], Problem 3. Given the chord (*subtensa*) of an arc less than half the circumference, and the chord of double the given arc, required to find the chord of the triple arc.[1]

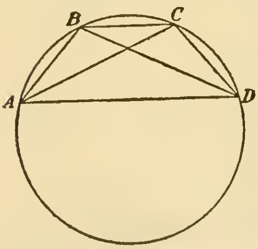
Solution ("rule").—Subtract the square of the chord of the given arc from the square of the chord of double the arc. The remainder divide by the chord of the given arc. The quotient will be the chord of the triple arc.[2]...

Fragment 2 [p. 50], Problem 6. Given the chord of an arc, find the chord of a third of the same arc.

Solution.—Take a third of the given chord; add something to it, and assuming the result to be the required chord compute the

---

[1] In modern notation: given 2 sin a and 2 sin $2a$, to find 2 sin $3a$.



[2] To prove this Pitiscus makes use of the fact that the chords of the three arcs form the sides and diagonals of an inscribed quadrilateral. If arc AB = arc BC = arc CD, $AB = BC = CB$ = chord of given arc. $AC = BD$ = chord of double the arc, and AD = chord of triple arc. According to a well-known theorem, $AC.BD = AB.CD + AD.BC$ or

$$\overline{AC}^2 = \overline{AB}^2 + AD.AB$$

Hence AD, or the chord of the triple arc, equals $\dfrac{\overline{AC}^2 - \overline{AB}^2}{AB}$.

Hence the proposition is proved.

given chord using the method of problem 3. Note the difference by plus or minus and, repeating the same operation on another assumed value of the required chord, mark the new difference by plus or minus. Having done this you will find the truth infallibly by the Rule of False.

Example.—Let the given arc *AD* or 30° be taken as 5176381. Required to find the chord of a third of the arc, namely, of the arc of 10°.

The given chord	= 5176381
One third of it	= 1725460
Increased value of the third	= 1730000
or	= 1740000
or	= 1750000

The first assumption is	= 1730000
The chord of the triple arc (30°) computed from this according to the method of problem 3	= 5138223
But it should be	= 5176381
Hence the difference is minus	= 38158

The

The second assumption is	= 1740000
The value of the chord of the triple arc computed by the method of problem 3	= 5167320
But it should be	= 5176381

Hence the difference is minus 9161

Now according to the Rule of False multiply across: that is, the first difference by the second assumption and the second difference by the first assumption. And since they are both negative subtract the products and you will have the number to be divided.

The first product	= 66394920000
The second product	= 15675530000

The number to be divided is 50719390000

Also from one of the minus numbers subtract the other and you will get the divisor.

One of them is	38158
The other is	9061

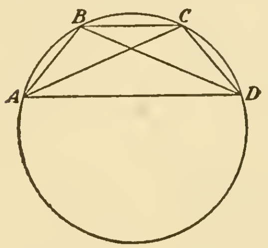
The divisor is 29097

The performed division will give for the chord AB the number 1743114. On this number perform an operation similar to that performed on each of the two assumed values, and again there will be a difference,—but very small, namely 3. Taking a number slightly greater than 1743114, namely the number 1743115, and repeating on it the above operation you will find that the chord AD will be almost equal to the given value 5176381 but in the end it will be a little greater. Therefore, the chord 1743115 will also be a little greater but nearer the truth than 1743114, as will appear from the computation; hence there will not be an appreciable difference between the given value of AD and the computed one.

Another Method by Algebra. Solution. Divide the given chord by $3x - x^3$.[1] The quotient will be the chord of a third of the given arc.

Proof of the Rule. The chord of any arc is equal to three roots less one cube, the root being equal to the chord of a third of this arc.[2]

This is demonstrated as follows: Let AD be the given chord of the arc $ABCD$. It is required to find the chord of AB, BC, or CD, [each of which is] a third of the arc.



Let x be the chord of the third of the arc. Hence each of the chords AC and BD of the double arcs will be $l4q - 1bq$,[3] as has been demonstrated in the solution of the preceding problem. Since $ABCD$ is an inscribed quadrilateral, the product of the diagonals AC and BD is equal to the sum of the products of the opposite sides, by proposition 54 of the first book [of the *Trigonometria*]. Multiply the diagonals and the square is $4x^2 - x^4$.[4]

Then multiply the side AB by the side CD,[5] that is x by x; x^2 is obtained. This, being subtracted from the square [made by]

---

[1] [Pitiscus uses l instead of x, and c for x^3. The chord of the lesser arc is equal to the root of the equation $3x - x^3 = AD$. In the translation we have used modern symbols.]

[2] [In modern notation, $2 \sin 3A = 3(2 \sin A) - (2 \sin A)^3$, which reduces to $\sin 3A = 3 \sin A - 4 \sin^3 A$.]

[3] The Pitiscus notation for $\sqrt{4x^2 - x^4}$.

[4] [Pitiscus here adds in parentheses the following characteristic remark: "Because to multiply a surd number by itself is nothing else than removing the sign l" i. e. the radical sign.]

[5] [Pitiscus retains the coefficient 1, writing $1l$ for $1x$, $1q$ for q or x^2, etc. In the translation this coefficient is omitted.]

the diagonals, that is, $4x^2 - x^4$ leaves $3x^2 - x^4$ for the rectangle (or the product) made by BC and AD. This rectangle $3x^2 - x^4$, being divided by BC, that is by x, will give as a result $AD = 3x - x^3$.[1]

Hence $3x - x^3$, where x is a chord of a third of the arc is equal to the chord of the given arc.

In consequence of the above, if the chord [k] of the given arc is equal to $3x - x^3$, the root [of the equation $3x - x^3 = k$] will be the chord of the third part.[2]

---

[1] [In modern notation:

$$AC.BD = AB.CD + AD.BC.$$

But

$$AC = BD = \sqrt{4x^2 - x^4},$$

by a previous demonstration

$$(x = AB = BC = CD.)$$

Hence

$$AC.BD = 4x^2 - x^4 = x^2 + x.AD.$$

$$\therefore AD = \frac{4x^2 - x^4 - x^2}{x} = 3x - x^3.]$$

[2] [This is equivalent, in modern symbols, to saying that

$$\sin 3A = 3 \sin A - 4 \sin^3 A.]$$

DE MOIVRE

On His Formula

(Translated from the Latin and from the French by Professor Raymond Clare Archibald, Brown University.)

De Moivre's Formula is usually stated in the form

$$(\cos x + i \sin x)^n = \cos nx + i \sin nx,$$

where n is any real number. The equivalent of this form was given by Euler (in 1748, Extract **E** below) and proved true for all real values of n (in 1749, Extract **F** below). The result is not explicitly stated in any of De Moivre's writings. But it will be observed that in more than one of them (1707–38, Extracts **A–D** below) the formula and its application were thoroughly familiar to him; and that in passages where it is suggested (1722, 1730) that certain eliminations shall be performed, on carrying these out we are led to exactly the formula associated with his name. This was made clear in Braunmühl's historical sketch in *Bibliotheca Mathematica*, series 3, vol. 2, p. 97–102, and in his *Vorlesungen über Geschichte der Trigonometrie*, part 2, 1903, p. 75–78. While Hutton's translations (*Philosophical Transactions*, abridged, vols. 5, 6, 8) have been the basis of the translations in Extracts **A**, **B**, and **D** they have not been slavishly followed as they were carefully compared with the originals. The original display of formulae and symbolism has been preserved except that in such an expression as $y.y$, y^2 has been substituted.

Abraham De Moivre was born in Champagne, France, in 1667, studied mathematics under Ozanam in Paris, and repaired in 1688 to London where he spent the remaining 66 years of his life. He was an intimate friend of Newton, and his notable mathematical publications led to his election not only as a member of the Royal Society, and of the Berlin Academy of Sciences, but also as a foreign associate of the Paris Academy of Sciences. Of his *Annuity Upon Lives* there were seven editions, five in English (1725, 1743, 1750, 1752, 1756), one in Italian (1776) much enlarged with notes of Gaeta and Fontana and the basis of lectures at the University of Pavia, and one in German (1906) by Czuber. Among his other publications, which display great analytic power, skill and invention, were *The Doctrine of Chances* (three editions, 1718, 1738, 1856), *Miscellanea Analytica* (1730) which brought about his election to the Berlin Academy, a number of papers in the *Philosophical Transactions*, and an important 8-page pamphlet of 1733 (*Approximatio ad summan terminorum binomii* $(a + b)^n$ *in seriem expansi*, English editions in the last two editions of *Doctrine of Chances*) presenting the first treatment of the probability integral and essentially of the normal curve. For a facsimile of the original edition of this pamphlet and for references to other discoveries of De Moivre, see *Isis*, vol. 8, 1926, p. 671–683. He was one of the commissioners appointed by the Royal Society in 1712 to arbitrate on the claims of Newton and Leibniz to the invention of the infinitesimal calculus.

Miss Clerke has recorded of De Moivre (*D.N.B.*) that he once said that he would rather have been Molière than Newton; and he knew his works and those of Rabelais almost by heart. In Pope's *Essay on Man* one finds (iii, l.103–104):

Who made the spider parallels design,
Sure as Demoivre, without rule or line?

A

"*Æquationum quarundam Potestatis tertiæ, quintæ, septimæ, novæ, & superiorum, ad infinitum usque pergendo, in terminis finitis, ad imstar Regularum pro Cubicis quae vocantur,* Cardani, *Resolutio Analytica,*" *Philosophical Transactions,* 1707, no. 309, vol. 25, p. 2368–2371.

"*The analytic solution of certain equations of the third, fifth, seventh, ninth and other higher uneven powers, by rules similar to those called Cardan's.*"

"The analytic solution of certain equations of the third, fifth, seventh, ninth and other higher uneven powers, by rules similar to those called Cardan's."

Let n denote any number whatever, y an unknown quantity or root of this equation, and a the absolute known quantity, or what is called the homogeneum comparationis[1]; let also the relation between these be expressed by the equation

$$ny + \frac{n^2-1}{2\times 3}ny^3 + \frac{n^2-1}{2\times 3}\times\frac{n^2-9}{4\times 5}ny^5 + \frac{(n^2-1)}{2\times 3}\times\frac{(n^2-9)}{4\times 5}\times\frac{(n^2-25)}{6\times 7}ny^7, \text{ etc.} = a.$$

From the nature of this series it is manifest, that if n be taken as any odd integer, either positive or negative, then the series will terminate and the equation become one of those above mentioned, the root of which is

$$(1)\qquad y = \tfrac{1}{2}\sqrt[n]{\sqrt{1+a^2}+a} - \frac{\frac{1}{2}}{\sqrt[n]{\sqrt{1+a^2}+a}}$$

$$\text{or } (2)\qquad y = \tfrac{1}{2}\sqrt[n]{\sqrt{1+a^2}+a} - \tfrac{1}{2}\sqrt[n]{\sqrt{1+a^2}-a}$$

$$\text{or } (3)\qquad y = \frac{\frac{1}{2}}{\sqrt[n]{\sqrt{1+a^2}-a}} - \tfrac{1}{2}\sqrt[n]{\sqrt{1+a^2}-a}$$

$$\text{or } (4)\qquad y = \frac{\frac{1}{2}}{\sqrt[n]{\sqrt{1+a^2}-a}} - \frac{\frac{1}{2}}{\sqrt[n]{\sqrt{1+a^2}+a}}\,[2]$$

---

[1] ["Homogeneum comparationis" in algebra was a name given by Vieta (1540–1603) to an equation's constant term, which he placed on the right-hand side of the equation and all the other terms on the left.]

[2] [In the denominator of the second term of the original, the second sign was − instead of +.]

For example, let it be required to find the root of the following equation of the fifth degree, $5y + 20y^3 + 16y^5 = 4$; in this case $n=5$ and $a=4$. According to (1) the root will then be

$$y = \tfrac{1}{2}\sqrt[5]{\sqrt{17}+4} - \frac{\tfrac{1}{2}}{\sqrt[5]{\sqrt{17}+4}}$$

whose numerical value is readily found. First $\sqrt{17}+4 = 8.1231$, whose logarithm is 0.9097164, of which the fifth part is 0.1819433, the number corresponding to which is $1.5203 = \sqrt[5]{\sqrt{17}+4}$. Also the arithmetic complement of 0.1819433 is 9.8180567, to which the number $0.6577 = \dfrac{1}{\sqrt[5]{\sqrt{17}+4}}$. Therefore the half difference of these numbers is $0.4313 = y$.

It may be here observed that, instead of the general root, it may be sufficient to take $y = \tfrac{1}{2}\sqrt[n]{2a} - \dfrac{\tfrac{1}{2}}{\sqrt[n]{2a}}$, whenever the number n is very large in comparison with unity. For example, if the equation were $5y + 20y^3 + 16y^5 = 682$; then $\log.2a = 3.1348143$, of which the fifth part, 0.6269628 corresponds to the number 4.236. Also its arithmetic complement is 9.3730372 which corresponds to the number 0.236. The half difference of these two numbers is $2=y$.

Again, if the terms of the preceding equation be alternately positive and negative, or which is the same thing, if the series be as follows:

$$ny + \frac{1-n^2}{2\times3}ny^3 + \frac{1-n^2}{2\times3}\times\frac{9-n^2}{4\times5}ny^5 + \frac{1-n^2}{2\times3}\times\frac{9-n^2}{4\times5}\times\frac{25-n^2}{6\times7}ny^7, \text{ etc.} = a,\text{[1]}$$

[1] [If $y=\sin\phi$, $a=\sin n\phi$ we have $\sin n\phi$ expressed in terms of $\sin\phi$, a result which Newton had already given in a letter of 13 June 1676 (*Commercium Epistolicum J. Collins et Aliorum* ed. Biot and Lefort, p. 106), and which De Moivre derived in an article, "A method of extracting the root of an infinite equation" in *Philosophical Transactions* for 1698, no. 240, vol. 20, 1699, p.190. This relation was derived as a special case of the following result with which the article opens: "If $az + bz^2 + cz^3 + dz^4 + ez^5$ etc. $= gy + hy^2 + iy^3 + ky^4 + ly^5$ etc. then will

$$zbe = \frac{g}{a}y + \frac{h - bA^2}{a}y^2 + \frac{i - 2bAB - cA^3}{a}y^3 + \frac{k - bB^2 - 2bAC - 3cA^2B - dA^4}{a}y^4 + \frac{l - 2bBC - 2bAD - 3cAB^2 - 3cA^2C - 4dA^3B - eA^5}{a}y^5 \text{ etc.''}$$

where A, B, C, D, etc.

its root will be

$$(1)\qquad y = \tfrac{1}{2}\sqrt[n]{a + \sqrt{a^2-1}} + \frac{\frac{1}{2}}{\sqrt[n]{a + \sqrt{a^2-1}}},$$

$$\text{or } (2)\qquad y = \tfrac{1}{2}\sqrt[n]{a + \sqrt{a^2-1}} + \tfrac{1}{2}\sqrt[n]{a - \sqrt{a^2-1}},$$

$$\text{or } (3)\qquad y = \frac{\frac{1}{2}}{\sqrt[n]{a - \sqrt{a^2-1}}} + \tfrac{1}{2}\sqrt[n]{a - \sqrt{a^2-1}},$$

$$\text{or } (4)\qquad y = \frac{\frac{1}{2}}{\sqrt[n]{a - \sqrt{a^2-1}}} + \frac{\frac{1}{2}}{\sqrt[n]{a + \sqrt{a^2-1}}}.$$

It should be here noted that if $\frac{n-1}{2}$ is an odd number, the sign of the root when found must be changed to the contrary sign.

If the equation $5y - 20y^3 + 16y^5 = 6$ be proposed, then $n=5$ and $a=6$. The root is equal to

$$\tfrac{1}{2}\sqrt[5]{6 + \sqrt{35}} + \frac{\frac{1}{2}}{\sqrt[5]{6 + \sqrt{35}}}.$$

Or, since $6 + \sqrt{35} = 11.916$ of which the logarithm is 1.0761304 and of which the fifth part is 0.2152561, the arithmetic complement being 9.7847439. Hence the numbers corresponding to these logarithms are 1.6415 and 0.6091 respectively whose half sum $1.1253 = y$.

But if it happen that a is less than unity then the second form of the root is rather to be preferred as more convenient for the purpose. Thus if the equation were

$$5y - 20y^3 + 16y^5 = \frac{61}{64}$$

$y = \frac{1}{2}\sqrt[5]{\frac{61}{64} + \sqrt{\frac{-375}{4096}}} + \frac{1}{2}\sqrt[5]{\frac{61}{64} - \sqrt{\frac{-375}{4096}}}$. And if by any

---

are respectively equal to the coefficients of y, y^2, y^3, y^4, etc., a result to which W. Jones refers (in his *Synopsis Palmariorum Matheseos*, London, 1706, p. 188) as a "*Theorem*" of "that Ingenious Mathematician *Mr. De Moivre*." In a review of this book of Jones in *Acta Eruditorum*, 1707, p. 176; this result is called "Theorema Moivræanum," a term assigned to a theorem which does not necessarily have any connection with trigonometric functions. The terms De Moivre's Formula, De Moivre's Theorem, applied to the formula we are considering, do not seem to have come into general use till the early part of the nineteenth century. Tropfke cites A. L. Crelle, *Lehrbuch der Elemente der Geometrie und der ebenen und sphärischen Trigonometrie*, Berlin, vol. 1, 1826, §335 for the use of the former term.]

means the fifth root of the binomial can be extracted the root will come out true and possible, although the expression seems to include an impossibility. Now the fifth root of the binomial $\frac{61}{64}+\sqrt{\frac{-375}{4096}}$ is $\frac{1}{4}+\frac{1}{4}\sqrt{-15}$, and of the binomial $\frac{61}{64}-\sqrt{\frac{-375}{4096}}$ is $\frac{1}{4}-\frac{1}{4}\sqrt{-15}$ whose semi-sum $\frac{1}{4}=$ y. But if the extraction can not be performed, or should seem too difficult, the thing may always be effected by a table of natural sines in the following manner.

To the radius 1 let $a=\frac{61}{64}=0.95112$ be the sine of a certain arc which therefore will be 72° 23′, the fifth part of which (because n = 5) is 14°28′; the sine of this is 0.24981, nearly $=\frac{1}{4}$. So also for equations of higher degree.[1]

B

"De Sectione Anguli," *Philosophical Transactions*, 1722, no. 374, vol. 32, p. 228–230.

"Concerning the section of an angle"

In the beginning of the year 1707, I fell upon a method by which a given equation of these forms

$$ny+\frac{n^2-1}{2\times 3}Ay^3+\frac{n^2-9}{4\times 5}By^5+\frac{n^2-25}{6\times 7}Cy^7,\text{ etc. }=a,$$

or

$$ny+\frac{1-n^2}{2\times 3}Ay^3+\frac{9-n^2}{4\times 5}By^5+\frac{25-n^2}{6\times 7}Cy^7,\text{ etc. }=a,$$

(where $A,B,C,\ldots$ represent the coefficients of the preceding terms) may have its roots determined in the following manner.

Set $a+\sqrt{a^2+1}=v$ in the first case and $a+\sqrt{a^2-1}=v$ in the second. Then will in the first case

$$y=\tfrac{1}{2}\sqrt[n]{v}-\frac{\frac{1}{2}}{\sqrt[n]{v}};\text{ and in the second } y=\tfrac{1}{2}\sqrt[n]{v}+\frac{\frac{1}{2}}{\sqrt[n]{v}}.$$

[1] [From this example it is clear that in 1707 De Moivre was in possession of the formula

$$\tfrac{1}{2}\sqrt[n]{\sin n\phi+\sqrt{-1}.\cos n\phi}+\tfrac{1}{2}\sqrt[n]{\sin n\phi-\sqrt{-1}.\cos n\phi}=\sin\phi,$$

where n is an odd integer. In 1730, as we shall presently see, De Moivre formulated a relation equivalent to the following:

$$\tfrac{1}{2}\sqrt[n]{\cos n\phi+\sqrt{-1}.\sin n\phi}+\tfrac{1}{2}\sqrt[n]{\cos n\phi-\sqrt{-1}.\sin n\phi}=\cos\phi,$$

where n is any positive integer.]

These solutions were inserted in the Philosophical Transactions, No. 309, for the months Jan. Feb. March of that year.

Now by what artifices these formulae were discovered will clearly appear from the demonstration of the following theorem.

In a unit circle let x denote the versed sine of any arc, and t that of another; and let the former arc be to the latter as 1 to n. Then, assuming two equations which may be regarded as known[1], namely

$$1 - 2z^n + z^{2n} = -2z^n t \text{ and } 1 - 2z + z^2 = -2zx,$$

on eliminating z, there will arise an equation by which the relation between x and t will be determined.

Corollary I.—If the latter arc be a semicircle, the equations will be

$$1 + z^n = 0,\ 1 - 2z + z^2 = -2zx,$$

from which if z be eliminated there will arise an equation by which will be determined the versed sines of the arcs which are to the semicircle taken once, or thrice, or five times, etc. as 1 to n.

Corollary II.—If the latter arc is a circumference, the equations are

$$1 - z^n = 0,\ 1 - 2z + z^2 = -2zx,$$

from which after z is eliminated will arise an equation by which are determined the versed sines of the arcs, which are to the circumference, taken once, twice, thrice, four times, etc., as 1 to n.

Corollary III.—If the latter arc is 60 degrees, the equations are

$$1 - z^n + z^{2n} = 0 \text{ and } 1 - 2z + z^2 = -2zx$$

from which on eliminating z, will arise an equation which determines the versed sines of the arcs which are to the arc of 60°, multiplied by 1,7,13,19,25, etc. or by 5,11,17,23,29,etc., as 1 to n.

If the latter arc be 120° the equations will be

$$1 + z^n + z^{2n} = 0 \text{ and } 1 - 2z + z^2 = -2zx,$$

---

[1] [Let x = versed sin $\phi = 1 - \cos\phi$, t = versed sin $n\phi = 1 - \cos n\phi$ then these equations are $1 - 2\cos n\phi z^n + z^{2n} = 0$ and $1 - 2\cos\phi z + z^2 = 0$. Compare quotation C, corollary I. The elimination of z gives

$$\sqrt[n]{\cos n\phi \pm \sqrt{\cos^2 n\phi - 1}} = \sqrt[n]{\cos n\phi \pm \sqrt{-1}.\sin n\phi}$$
$$= \cos\phi + \sqrt{-1}.\sin n\phi,$$

or $(\cos\phi + \sqrt{-1}.\sin\phi)^n = \cos n\phi + \sqrt{-1}.\sin n\phi$, De Moivre's formula, for n an odd integer.]

from which if z be eliminated there will arise an equation, by which are determined the versed sines of the arcs, which are to the arcs of 120°, multiplied by 1,4,7,10,13, etc., or by 2,5,8,11,14, etc. as 1 to n.

C

[A. De Moivre], *Miscellanea Analytica*, London, 1730, pp. 1–2.

Lemma 1.—*If* l *and* x *are the cosines of two arcs* A *and* B *of a circle of radius unity, and if the first arc is to the second as the number* n *is to unity then*

$$x = \tfrac{1}{2}\sqrt[n]{l + \sqrt{l^2-1}} + \frac{\frac{1}{2}}{\sqrt[n]{l + \sqrt{l^2-1}}}$$

Corollary I.—Set $\sqrt[n]{l + \sqrt{l^2-1}} = z$; then will $z^n = l + \sqrt{l^2-1}$ or $z^n - l = \sqrt{l^2-1}$, or, squaring both sides $z^{2n} - 2lz^n + l^2 = l^2 - 1$. Cancelling equal terms on each side and having made the proper transposition $z^{2n} - 2lz^n + 1 = 0$. From what was assumed $\sqrt[n]{l + \sqrt{l^2-1}} = z$, it follows from the above lemma that

$$x = \tfrac{1}{2}z + \frac{\frac{1}{2}}{z}, \text{ or } z^2 - 2xz + 1 = 0.$$

Corollary II.—If between the two equations $1 - 2lz^n + z^{2n} = 0$, $1 - 2xz + z^2 = 0$, the quantity z be eliminated there will arise a new equation defining a relation between the cosines l and x, providing the arc A is less than a quadrant.

Corollary III.—But if the arc A is greater than a quadrant then its cosine will be $-l$, from which it results that the equations will turn out to be $1 + 2lz^n + z^{2n} = 0$, $1 - 2xz + z^2 = 0$; and if z be eliminated between these there will arise a new equation expressing the relation between the cosines l and x.

Corollary IV.—And in particular if z be eliminated between the equations $1 \mp 2lz^n + z^{2n} = 0$, $1 - 2xz + z^2 = 0$ there will arise a new equation expressing a relation between the cosine of the arc A (less or greater than a quadrant according as l has the negative or positive sign) and all the cosines of the arcs $\frac{A}{n}$, $\frac{C-A}{n}$, $\frac{C+A}{n}$, $\frac{2C-A}{n}$, $\frac{2C+A}{n}$, $\frac{3C-A}{n}$, $\frac{3C+A}{n}$, etc., in which series of arcs C denotes the entire circumference.[1]

---

[1] [That is $\sqrt[n]{\cos A \pm i \sin A} = \cos \frac{2k\pi \pm A}{n} + i \sin \frac{2k\pi \pm A}{n}$, $k = 0,1,2,3, \ldots$.]

D

"*De Reductione* Radicalium *ad simpliciores terminos, seu de extrahenda* radice *quacunque data ex* Binomio $a + \sqrt{+b}$, *vel* $a + \sqrt{-b}$." Epistola, *Philosophical Transactions*, 1739, no. 451, vol. 40, p. 463–478.

"*On the reduction* of radicals to simpler terms, or the extraction of roots *of any* binomial $a + \sqrt{+b}$ or $a + \sqrt{-b}$. Letter."

[The paper consists almost wholly in the discussion of four problems. Our quotation is of problems 2–3, p. 472–74.]

Problem II.—*To extract the cube root of the impossible binomial* $a + \sqrt{-b}$.

Suppose that root to be $x + \sqrt{-y}$, the cube of which is $x^3 + 3x^2\sqrt{-y} - 3xy - y\sqrt{-y}$. Now put $x^3 - 3xy = a$, and $3x^2\sqrt{-y} - y\sqrt{-y} = \sqrt{-b}$. Then the squares of these will give two new equations, namely

$$x^6 - 6x^4y + 9x^2y^2 = a^2$$
$$-9x^4v + 6x^2y^2 - y^3 = -b.$$

Then the difference of these squares is

$$x^6 + 3x^4y + 3x^2y^2 + y^3 = a^2 + b;$$

the cube root of which is $x^2 + y = \sqrt[3]{a^2 + b} = m$, say. Hence $x^2 + y = m$, or $y = m - x^2$ which value of y substituted in the equation $x^3 - 3xy = a$ gives $x^3 - 3mx + 3x^3 = a$, or $4x^3 - 3mx = a$; which is the very same equation as has been before deduced from the equation $2x = \sqrt[3]{a + \sqrt{-b}} + \sqrt[3]{a - \sqrt{-b}}$. Nevertheless it does not follow that in the equation $4x^3 - 3mx = a$, the value of x can be found by the former equation since it consists of two parts each including the imaginary quantity $\sqrt{-b}$; but this will best be done by means of a table of sines.

Therefore let the cube root be extracted of the binomial $81 + \sqrt{-2700}$. Put $a = 81$, $b = 2700$; then $a^2 + b = 6561 + 2700 = 9261$, the cube root of which is 21, which set equal to m makes $3mx = 63x$. Hence the equation to be solved will be $4x^3 - 63x = 81$, which being compared with the equation for the cosines, namely $4x^3 - 3r^2x = r^2c$[1] gives $r^2 = 21$, hence $r = \sqrt{21}$ and there-ore $c = \frac{a}{r^2} = \frac{81}{21} = \frac{27}{7}$.

---

[1] [If this equation be put in the form $4\left(\frac{x}{r}\right)^3 - 3\left(\frac{x}{r}\right) = \frac{c}{r}$ it may be regarded as equivalent to the trigonometric formula $4\cos^3\frac{A}{3} - 3\cos\frac{A}{3} = \cos A$, if

To find the circular arc corresponding to the radius $\sqrt{21}$ and $c=\frac{27}{7}$, put the whole circumference equal to C, and take the arcs $\frac{A}{3}$, $\frac{C-A}{3}$, $\frac{C+A}{3}$, which will easily be known by a trigonometrical calculation, especially by using logarithms; then the cosines of the arcs to the radius $\sqrt{21}$ will be the three roots of the quantity x; since $y = m - x^2$, there will therefore be as many values of y, and thence a triple value of the cube root of the binomial $81 + \sqrt{-2700}$, which must now be accommodated to numbers.[1]

Make then $\sqrt{21}$ to $\frac{27}{7}$ as the tabular radius is to the cosine of an arc A, which will be nearly 32°42′. Hence the arc $C-A$ will be 327°18′, and $C+A$ 392°42′, of which the third parts will be 10°54′, and 109°6′, and 130°54′. But now as the first of these is less than a quadrant, its cosine, that is, the sine of 79°6′ ought to be considered as positive; and both the other two being greater than a quadrant, their cosines, that is the sines of the arcs 19°6′ and 40°54′, must be considered as negative. Now by trigonometrical calcu-

---

$\frac{c}{r} = \cos A$ and $\frac{x}{r} = \cos\frac{A}{3}$. De Moivre identifies the problem of finding the cube root with that of trisecting an angle.]

[1] [In general terms the argument is as follows:

$c=r\cos A$, $x=r\cos\frac{A}{3}$ and by comparing the two cubic equations $m=r^2$, $a=r^2c$. Therefore $r=\sqrt{m}$, $c=\frac{a}{m}=\sqrt{m}\cos A$. Then $x=\sqrt{m}\cos\frac{A}{3}$,

$$x=\sqrt{m}\cos\frac{C-A}{3},$$

$x = \sqrt{m}\cos\frac{C+A}{3}$ are the three roots of the cubic equation. But since $a=r^2c=\sqrt{m^3}\cos A$, and $b=m^3-a^2=m^3(1-\cos^2 A)=m^3\sin^2 A$, $x=\sqrt{m}\cos\frac{A}{3}$, $\ldots$, $y=m-x^2=m\left(1-\cos^2\frac{A}{3}\right)=m\sin\frac{A}{3}\ldots$ Hence on substituting in the equation $\sqrt{a+i\sqrt{b}}=x+i\sqrt{y}$ we get

$$[\sqrt{m^3}(\cos A+i\sin A)]^{\frac{1}{3}} = \sqrt{m}\left\{\cos\frac{A}{3} + i\sin\frac{A}{3}\right\}.$$

or

$$\sqrt{m}\left(\cos\frac{C-A}{3} + i\sin\frac{C-A}{3}\right),$$

or

$$\sqrt{m}\left(\cos\frac{C+A}{3} + i\sin\frac{C+A}{3}\right).]$$

lation it appears, that these sines, to radius $\sqrt{21}$ will be 4.04999 and -1.4999, and -3.0000, or $\frac{9}{2}$, and $-\frac{3}{2}$ and -3. Hence there will be as many values of the quantity y, namely all those represented by $m-x^2$, namely $21-\frac{81}{4}$, and $21-\frac{9}{4}$, and $21-9$, that is $\frac{3}{4}$, $\frac{75}{4}$, 12 and the square roots of which are $\frac{1}{2}\sqrt{3}$, $\frac{5}{2}\sqrt{3}$, $2\sqrt{3}$. Therefore the values of $\sqrt{-y}$ will be $\frac{1}{2}\sqrt{-3}$, $\frac{5}{2}\sqrt{-3}$, $2\sqrt{-3}$. Hence the values of $\sqrt[3]{81+\sqrt{-2700}}$ are $\frac{9}{2}+\frac{1}{2}\sqrt{-3}$, $-\frac{3}{2}+\frac{5}{2}\sqrt{-3}$, and $-3+\frac{1}{2}\sqrt{-3}$. And by proceeding in the same manner, there will be found the three values of $\sqrt[3]{81-\sqrt{-2700}}$, which are $\frac{9}{2}-\frac{1}{2}\sqrt{-3}$, $\frac{3}{2}-\frac{5}{2}\sqrt{3}$, and $-3-\frac{1}{2}\sqrt{-3}$.

There have been several authors, and among them the eminent Wallis, who have thought that those cubic equations which are referred to the circle, may be solved by the extraction of the cube root of an imaginary quantity, and of $81+\sqrt{-2700}$, without regard to the table of sines, but that is a mere fiction and a begging of the question. For on attempting it, the result always recurs back again to the same question as that first proposed. And the thing cannot be done directly, without the help of the table of sines, especially when the roots are irrational, as has been observed by many others.

Problem III.—*To extract the nth root of the impossible binomial* $a+\sqrt{-b}$.

Let that root be $x+\sqrt{-y}$; then making $\sqrt[n]{a^2+b}=m$, and $\frac{n-1}{n}$[1] $=p$, n any integer, describe, or conceive to be described, a circle, the radius of which is $\sqrt{m}$, in which take an arc A the cosine of which is $\frac{a}{m^p}$, and let C be the whole circumference. To the same radius take the cosines of the arcs $\frac{A}{n}$, $\frac{C-A}{n}$, $\frac{C+A}{n}$, $\frac{2C-A}{n}$, $\frac{2C+A}{n}$, $\frac{3C-A}{n}$, $\frac{3C+A}{n}$, etc., till the number of them be

---

[1] [This should be $\frac{n-1}{2}=p$.]

equal to n. Then all these cosines will be so many values of x; and the quantity y will always be $m-x^2$.[1]

E

Euler, *Introductio in Analysin Infinitorum*, Lausanne, 1748, vol.1, Chapter 8, "De quantitatibus transcendentibus ex circulo ortis" ["On transcendental quantities derived from the circle,"] p.97–98, §§132–133; Reprinted in *Leonhardi Euleri Opera Omnia*, Leipzig, series 1, vol.8, 1922, p.140–141.

132. Since $(\sin.z)^2 + (\cos.z)^2 = 1$, on decomposing into factors we get $(\cos.z + \sqrt{-1}.\sin.z)(\cos.z - \sqrt{-1}\ \sin.z) = 1$. These factors, although imaginary are of great use in the combination and multiplication of arcs. For example, let us seek the product of these factors

$$(\cos.z + \sqrt{-1}.\sin.z)(\cos.y + \sqrt{-1}.\sin.y),$$

we find

$$\cos.y\ \cos.z - \sin.y\ \sin.z + \sqrt{-1}.(\cos.y\ \sin.z + \sin.y\ \cos.z);$$

but since

$$\cos.y\ \cos.z - \sin.y\ \sin.z = \cos.(y+z)$$

and

$$\cos.y\ \sin.z + \sin.y\ \cos.z = \sin.(y+z)$$

we obtain the product

$$(\cos.y + \sqrt{-1}.\sin.y)(\cos.z - \sqrt{-1}.\sin.z) = \cos.(y+z) + \sqrt{-1}.\sin.(y+z).$$

Similarly

$$(\cos.y - \sqrt{-1}.\sin.y)(\cos.z - \sqrt{-1}.\sin.z) = \cos.(y+z) - \sqrt{-1}.\sin.(y+z).$$

In the same way

$$(\cos.x \pm \sqrt{-1}.\sin.x)(\cos.y \pm \sqrt{-1}.\sin.y)(\cos.z + \sqrt{-1}.\sin.z) = \cos.(x+y+z) \pm \sqrt{-1}.\sin.(x+y+z).$$

---

[1] [That is,—

$$(\sqrt{m^n}(\cos A + i\ \sin A))^{\frac{1}{n}} = \sqrt{m}\left(\cos\frac{2k\pi \pm A}{n} + i \sin \frac{2k\pi \pm A}{n}\right)$$
$$k = 0,1,2,\ldots\frac{n-1}{2} \text{ if } n \text{ is odd; or } k=1,2,\ldots\frac{n}{2} \text{ if } n \text{ is even,}$$

De Moivre's theorem for any unit fraction.

Practically all of extract D is given by De Moivre in a communication dated April 29, 1740, in N. Saunderson, *The Elements of Algebra*, Cambridge, vol. 2, 1740, p.744–748.]

133. Hence it follows that

$$(\cos.z \pm \sqrt{-1}.\sin.z)^2 = \cos.2z \pm \sqrt{-1}.\sin.2z,$$

and

$$(\cos.z \pm \sqrt{-1}.\sin.z)^3 = \cos.3z \pm \sqrt{-1}.\sin.3z;$$

and in general

$$(\cos.z \pm \sqrt{-1}.\sin.z)^n = \cos.nz \pm \sqrt{-1}.\sin.nz.$$

From these, by virtue of the double signs, we deduce

$$\cos.nz = \frac{(\cos.z + \sqrt{-1}.\sin.z)^n + (\cos.z - \sqrt{-1}.\sin.z)^n}{2}$$

and

$$\sin.nz = \frac{(\cos.z + \sqrt{-1}.\sin.z)^n - (\cos.z - \sqrt{-1}.\sin.z)^n}{2}$$

Developing these binomials into series we get

$$\cos.nz = (\cos.z)^n - \frac{n(n-1)}{1.2}(\cos.z)^{n-2}(\sin.z)^2 + \frac{n(n-1)(n-2)(n-3)}{1.2.3.4}(\cos.z)^{n-4}(\sin.z)^4 - \frac{n(n-1)(n-2)(n-3)(n-4)(n-5)}{1.2.3.4.5.6}(\cos.z^{n-6})(\sin.z)^6 + \text{etc.}$$

and

$$\sin.nz = \frac{n}{1}(\cos.z)^{n-1}\sin.z - \frac{n(n-1)(n-2)}{1.2.3}(\cos.z^{n-3})(\sin.z)^3 + \frac{n(n-1)(n-2)(n-3)(n-4)}{1.2.3.4.5}(\cos.z)^{n-5}(\sin.z)^5 - \text{etc.}$$ [1]

---

[1] [A little later in the chapter Euler gives the formula;

$$\cos.v = \frac{e^{+v\sqrt{-1}} + e^{-v\sqrt{-1}}}{2}, \quad \sin.v = \frac{e^{+v\sqrt{-1}} - e^{-v\sqrt{-1}}}{2\sqrt{-1}};$$

and

$$e^{+v\sqrt{-1}} = \cos.v + \sqrt{-1}.\sin.v,$$
$$e^{-v\sqrt{-1}} = \cos.v - \sqrt{-1}.\sin.v.$$

Roger Cotes gave much earlier the equivalent of the formula

$$\log(\cos x + i \sin x) = i x$$

(*Philosophical Transactions*, 1714, vol. 29, 1717, p.32) under the form: Si quadrantis circuli quilibet arcus, radio CE descriptus, sinum habeat CX, sinumque complementi ad quadrantem XE: sumendo radium CE, pro Modulo, arcus erit rationis inter EX + XC$\sqrt{-1}$ & CE mensura ducta in $\sqrt{-1}$.]

F

Euler, "Recherches sur les racines imaginaires des equations", *Histoire de l'Academie Royale des Sciences et Belles Lettres*, Berlin, vol.5 (1749), 1751, p.222–288. The following extract covers §79–85, p.265–268.

§79. Problem I.—*An imaginary quantity being raised to a power which is any real number, determine the form of the imaginary which results.*

Solution. Suppose $a + b\sqrt{-1}$ is the imaginary quantity, and m the real exponent; it is required to determine M and N, such that

$$(a + b\sqrt{-1})^m = \mathrm{M} + \mathrm{N}\sqrt{-1}.$$

Set $\sqrt{(a^2 + b^2)} = c$; c will always be a real positive quantity since we do not here regard the ambiguity of the sign $\sqrt{\ }$. Further, let us seek the angle ϕ such that its sine is equal to $\frac{b}{c}$ and cosine $\frac{a}{c}$, here having regard to the nature of the quantities a and b if they are positive or negative. It is certain that one can always find this angle ϕ whatever the quantities a and b are, provided that they are real, as we suppose. Now having found this angle ϕ which will be always real, one will at the same time find other angles whose sine is $\frac{b}{c}$ and cosine $\frac{a}{c}$ are the same; namely, on setting π for the angle of 180°, all the angles ϕ, $2\pi+\phi$, $4\pi+\phi$, $6\pi+\phi$, $8\pi+\phi$, etc. to which one may add $-2\pi+\phi$, $-4\pi+\phi$, $-6\pi+\phi$, $-8\pi+\phi$, etc. That being said $a + b\sqrt{-1} = c(\cos\phi + \sqrt{-1}.\sin\phi)$ and raising to the proposed power

$$(a + b\sqrt{-1})^m = c^m(\cos\phi + \sqrt{-1}.\sin\phi)^m$$

where c^m will always have a real positive value. In consequence of the demonstration that

$$(\cos\phi + \sqrt{-1}.\sin\phi)^m = \cos m\phi + \sqrt{-1}.\sin m\phi,$$

where it is to be remarked that since m is a real quantity, the angle $m\phi$ will be also real and hence also its sine and cosine, we will have

$$(a + b\sqrt{-1})^m = c^m(\cos m\phi+\sqrt{-1}.\sin m\phi).$$

Or, the power $(a + b\sqrt{-1})^m$ is contained in the form

$$\mathrm{M} + \mathrm{N}\sqrt{-1},$$

on setting $M=c^m \cos m\phi$ and $N=c^m \sin m\phi$ where $c = \sqrt{(a^2+b^2)}$, and $\cos\phi=\frac{a}{c}$ and $\sin\phi=\frac{b}{c}$. C.QF.T.[1]

§80. *Corollary* I.—In the same way that $(\cos\phi + \sqrt{-1}.\sin\phi)^m = \cos m\phi+\sqrt{-1}.\sin m\phi$, is also $(\cos\phi - \sqrt{-1}.\sin\phi)^m = \cos m\phi - \sqrt{-1}.\sin m\phi$; and hence

$$(a - b\sqrt{-1})^m = c^m (\cos m\phi-\sqrt{-1}.\sin m\phi),$$

whereϕ is the same angle as in the preceding case.

§81. *Corollary* II.—If the exponent m is negative, since $\sin -m\phi = -\sin m\phi$ and $\cos -m\phi$ - $\cos m\phi$, then

$$(\cos\phi \pm \sqrt{-1}.\ \sin\phi)^{-m} = \cos m\phi \mp \sqrt{-1}.\sin m\phi$$

and

$$(a \pm b\sqrt{-1})^{-m} = c^{-m}(\cos m\phi \mp \sqrt{-1}.\sin m\phi).$$

§82. *Corollary* III.—If m is an integer positive or negative the formula $(a + b\sqrt{-}\ 1)^m$ has only a single value; for whatever is substituted for ϕ of all the angles $\pm 2\pi + \phi$, $\pm 4\pi + \phi$, $\pm 6\pi + \phi$, etc., one always finds the same values for $\sin m\phi$ and $\cos m\phi$.

§83. *Corollary* IV.—But if the exponent m is a rational number $\frac{\mu}{\nu}$, the expression $(a + b\sqrt{-1})^{\frac{\mu}{\nu}}$ will have as many different values as there are units in ν. For, on substituting for ϕ the angle one will obtain as many different values for $\sin m\phi$ and $\cos m\phi$ as the number ν contains of units.

§84. *Corollary* V.—Whence it is clear that if m is an irrational number, or incommensurable to unity, the expression $(a + b\sqrt{-1})^m$ will have an infinite number of different values, since all the angles ϕ, $\pm 2\pi + \phi$, $\pm 4\pi+\phi$, $\pm 6\pi+\phi$, etc., will furnish different values for $\sin m\phi$ and $\cos m\phi$.

§85. *Scholium* I.—The foundation of the solution of this problem is that $(\cos\phi + \sqrt{-1}.\sin\phi)^m = \cos m\phi+\sqrt{-1}\ \sin m\phi$, whose truth is proved by known theorems regarding the multiplication of angles. For having two angles ϕ and θ,

$$(\cos\phi+\sqrt{-1}.\sin\phi)(\cos\theta+\sqrt{-1}.\sin\theta) = \cos\ (\phi + \theta) + \sqrt{-1}.\ \sin(\phi+\theta),$$

which is clear by actual multiplication which gives $\cos\phi \cos\theta - \sin\phi \sin\theta + (\cos\phi.\sin\theta + \sin\phi.\cos\theta)\sqrt{-1}$. But

$\cos\phi.\cos\theta - \sin\phi.\sin\theta = \cos\ (\phi + \theta)$, and $\cos\phi \sin\theta + \sin\phi \cos\theta = \sin\ (\phi + \theta)$.

[1] [Ce qu'il fallait trouver, which was to be found.]

Hence one may readily deduce the consequence that

$$(\cos\phi + \sqrt{-1}.\sin\phi)^m = \cos m\phi + \sqrt{-1}.\sin m\phi,$$

when the exponent m is an integer. All doubt that the same formula is also true, when m is any number, is removed by differentiation, after having taken logarithms. For, taking logarithms, there results

$$m\mathrm{l}(\cos\phi + \sqrt{-1}.\sin\phi) = \mathrm{l}(\cos m\phi + \sqrt{-1}\sin m\phi).$$

Treating the angle ϕ as a variable quantity we will have

$$\frac{-m\, d\phi\, \sin\phi + m\, d\phi\sqrt{-1}.\cos\phi}{\cos\phi + \sqrt{-1}.\sin\phi} = \frac{-m\, d\phi\, \sin m\phi + m\, d\phi\sqrt{-1}.\cos m\phi}{\cos m\phi + \sqrt{-1}.\sin m\phi}.$$

On multiplying the numerators by $-\sqrt{-1}$, one obtains

$$\frac{m\, d\phi(\cos\phi + \sqrt{-1}\,.\,\sin\phi)}{\cos\phi + \sqrt{-1}.\sin\phi} = \frac{m\, d\phi(\cos m\phi + \sqrt{-1}.\sin m\phi)}{\cos m\phi + \sqrt{-1}.\sin m\phi} = m\, d\phi,$$

which is an identical equation.[1]

[1] [This proves that either $(\cos\phi + \sqrt{-1}.\sin\phi)^m$ and $\cos m\phi + \sqrt{-1}.\sin m\phi$ are equal for all values of ϕ or differ by a constant. For $\phi = 0$ they are equal. Hence the proof is completed.]

CLAVIUS AND PITISCUS

On Prosthaphaeresis

(Translated from the Latin by Professor Jekuthiel Ginsburg, Yeshiva College, New York City.)

In the years immediately preceding the discovery of logarithms, mathematicians made use of a method called *prosthaphaeresis* to replace the operations of multiplication and division by addition and subtraction. The method was based on the equivalent of the formula

$$\cos(A - B) - \cos(A + B) = 2 \sin A \sin B$$

Nicolaus Raymarus Ursus Dithmarsus used it in the solution of spherical triangles where it is necessary to find the fourth proportional to the sinus totus (radius), sin A, and sin B. Christopher Clavius (1537–1612) extended the method to the cases of secants and tangents; in fact he showed how to find the product of any two numbers by this method, thus in a way anticipating the theory of logarithms.

In the first fragment translated below, Clavius shows how the product of two sines may be found by the method of prostaphaeresis. This seems to be the original essence of the discovery as presented by Raymarus Ursus. It is applicable only when each of the two factors is less that the sinus totus and may therefore be considered as the sine of some arc.

In the second fragment Clavius shows how to proceed in the case when one of the numbers to be multiplied is greater than the sinus totus.

Both fragments are from his *Astrolabium* (Rome, 1593), Book I, lemma 53.

Fragment I

Let for example it be required to find the declination[1] of 17° 45′π.

Since it is true that the sinus totus is to the sine of the maximum declination[2] of the point as the sine of the distance of the given

---

[1] [The declination is the distance from a point on the celestial sphere to the equator. (The distance of course is to be measured on the arc of the meridian.)

The point 17°45′π is a point in the third sign of the zodiac. Since every sign of the zodiac is 30° and since the first sign begins at the intersection of the ecliptic with the equator, the distance of the point from that intersection equals $30° + 30° + 17°45' = 77°45'$.]

[2] [The angle formed by the equator and ecliptic.]

ecliptical point from the nearest equinoctial point to the sine of the declination of the same point.[1]

Hence, by prosthaphaeresis,

Max. declination	23°30′	Complement of greater	12°15′
Distance from equinox	77°45′	Smaller Arc	23°30′

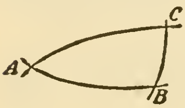
Sum of complement and smaller arc 35°45′. The sine is 3842497
Diff. of complement and smaller arc 11°15′. The sine is 1950903

The sum of the sines 7793400
Half of the sum 3896700 declination 22°56′[2]

Fragment II

When ratio of the sinus totus to a number less than itself is equal to the ratio of a number greater than the sinus totus to the required number[3] proceed as follows: The third number, which is greater than the sinus totus, should be divided by the sinus totus. The quotient will be the number obtained when seven

[1] [Let AB be the arc of the equator, AC an arc of the ecliptic, A the nearest point of the equinox,—that is, the intersection of the equator with the ecliptic,



—C a point in the third sign of the ecliptic whose distance AC from the equinox is equal to 2 signs plus 17°45′, or 77°45′, CB the distance of the point C from the equator measured on the arc of the meridian. We see that CB is perpendicular to AB and triangle ABC is a right spherical triangle. We have then

$$\sin B : \sin A = \sin AC : \sin CB$$

where

$$\sin B = \sin 90° = 1 \text{ (sinus totus)},\ A = 23°30'$$

(this being the known angle of intersection of the ecliptic and equator). Sin $AC = \sin 77°\,45'$, and CB is the arc to be computed; hence in modern nota-

$$1 : \sin 23°30' = \sin 77°45' : \sin CB,$$

tion which is adopted for the purposes of prosthaphaeresis.]

[2] [Explanation. According to a theorem proved by Clavius on p. 179,

$$1 : \sin A = \sin B : \tfrac{1}{2}\,[\sin(\overline{90° - A} + B) - \sin(\overline{90° - A} - B)]$$

where

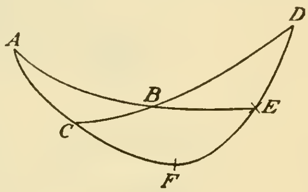
$A = 77°45'$, $90° - A = 12°15'$, $B = 23°30'$; $\overline{90° - A} + B = = 23°30' + 12°15' = 35°45'$, $\overline{90° - A} - B = 23°30' - 12°15' = = 11°15'$. Sin $(\overline{90° - A} + B)$ and sin $(\overline{90° - A} - B)$ are to be added or subtracted according as $90° - A$ is greater or less than B. Hence the required product of sin 23° 30′ by sin 77°45′ = ½ sin 35°45′ − ½ sin 11°15′.]

[3] [*I.e.* 10,000,000: $A = B{:}x$ where $A < 10^7$ and $B > 10^7$. The difficulty here is due to the fact that B cannot be represented as a sine.]

figures are cut off on the right (ad dexterum). These seven figures form the remainder. Then the proportion "the sinus totus is to the smaller number as the residue is to the required number" will be adapted to the use of prosthaphaeresis if the arcs of the smaller number and the residue considered as sines are taken from the table. To the fourth proportional thus found should be added the product of the smaller number by the quotient of the above division.[1]

Pitiscus on Prosthaphaeresis[2]

Problem I.—Given a proportion in which three terms are known. To solve the proportion in which the first term is the radius, while the second and third terms are sines, avoiding multiplication and division.[3]



Find the sum of the complements of the arcs corresponding to these and you will have a right spherical triangle which fits in the fourth case of spherical triangles which is solvable by prosthaphaeresis alone.

For example, given the proportion:

"radius AE is to sin EF as sin AB is to sin BC."

---

[1] [As an illustration, Clavius considers the product of 3,912,247 by 11,917,537 or the proportion 10^7: 391,247 = 11,917,535: x, in which the third term, 11,917,535, is greater than 10^7. Dividing it by 10^7 the quotient is 1 and the remainder 1,917,535. Clavius then considers the auxiliary proportion 10^7: 391,247 = 1,917,535: x. From the tables he finds that 391,247 = sin 23°2′, 1,917,535 = sin 11°3′. Hence the proportion becomes

$$10^7 : \sin 23°2' = \sin 11°3' : x.$$

He proceeds then as above

$$10^7 : \sin 23°2' = \sin 11°3' : \tfrac{1}{2}[\sin (90° - 23°2' + 11°3') - \sin (90° - 23°2' - 11° 3')]$$

which is equivalent to

$$10^7 : \sin 23°2' = \sin 11°3' : \tfrac{1}{2}[\sin 78°1' - \sin 55°55'],$$

which gives

$$x = 749{,}923.$$

To this result should be added the product of the second term 391,247 by the quotient 1.]

. . . .

[2] *Trigonometria*, 1612 ed., p. 149.

[3] [*I.e.*, solve 1: sin A = sin B:x, for x, without using either multiplication or division.]

For the given arcs EF, AB in the second and third place find the complements ED, BE and you will have a triangle BED which is right-angled at E in which the required quantity BC is the complement of the side DB. This you will find by the fourth axiom of the spherical triangles without multiplication or division.

Let the side $AB = 42°$, and

$$EF = 48°25',$$

Then

$$BE = 48°,$$
$$DE = 41°35',$$

from which it will follow

	$DE = 41°35' = 41°35'$		
	$BE = 48°0'$ compl. $42°0'$		
89°35′	83°35′	sin	9937354
0°25′		sin	72721
			10010075
[Half of this is]			5005037

which is the sine of the required arc BC, or $36°2'$.

CLAVIUS

On Prosthaphaeresis as Applied to Trigonometry

(Translated from the Latin by Professor Jekuthiel Ginsburg, Yeshiva College, New York City.)

The following fragment is from Christopher Clavius, *Astrolabium*, (Rome, 1593), pp. 179–180. It contains the proof of the formula

$$\cos (A - B) - \cos (A + B) = 2 \sin A, \sin B.$$

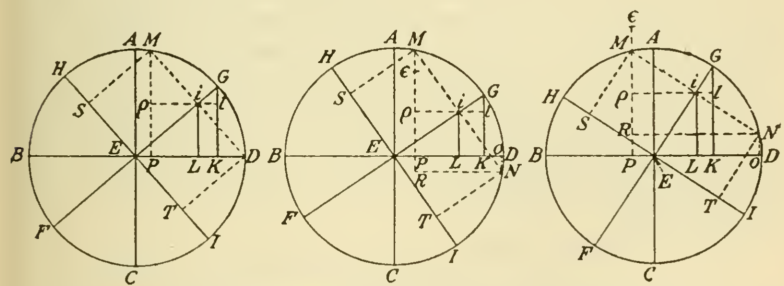
Clavius considers three cases:

$$1) \ A + B = 90^\circ,$$
$$2) \ A + B < 90^\circ,$$
$$3) \ A + B > 90^\circ.$$

In the first case the formula becomes

$$\sin 2A = 2 \sin A \cos A.$$

Clavius credits Nicolaus Raymarus Ursus Dithmarsus with the discovery of the theorem, but according to A. Braunmühl (*Vorlesungen über Geschichte der Trigonometrie* p. 173) the latter proved only two of the cases, namely the second and the third. In any case Clavius was the first to publish the theorem and the proof in the complete form.



Clavius, as may be seen from his opening note, used the theorem as an introduction to the method of "prosthaphaeresis" which, in the years immediately preceding Napier's discovery of logarithms, was used by mathematicians like Raymarus Ursus, Bürgi, Clavius, and others to replace multiplication and division by the operation of addition and subtraction. The bearing of the subject upon the theory, if not the invention, of logarithms is apparent. The translation is as follows:

Lemma LIII.—Three or four years ago Nicolaus Raymarus Ursus Dithmarsus published a leaflet in which he proposed, among other things, an ingenious device by means of which he solved many spherical triangles by prosthaphaeresis[1] only. But since it is usable only when the sines are assumed in a proportion and when the sinus totus takes the first place, we will attempt here to make the doctrine more general, so that it will hold not only for sines, and [not only] when the sinus totus is in the first place, but also for tangents, secants, versed sines, and other numbers, no matter whether the sinus totus appears at the beginning or in the middle, and even when it does not appear at all. These things are entirely new and full of pleasure and satisfaction [iucunditatis ac voluptatis plena].

THEOREM.—The sinus totus is to the sine of any arc as the sine of another arbitrary arc [sinus alterius cujuspiam arcus] is to a quantity composed of these two arcs in a way required for the purpose of prosthaphaeresis. The smaller is to be added to the complement of the greater and the sine of the sum is to be taken.[2] Then

1. If the minor arc is equal to the complement of the greater (that is, when the two arcs are together equal to a quadrant), half of the computed sine will be the required fourth term of the proportion.[3]

2. If, however, the smaller arc is less than the complement of the greater (which happens when the sum of the two arcs is less than a quadrant of a circle), the smaller arc is subtracted from the complement of the greater so that we now have the difference between the same arcs that have been added before, and the sine of this difference[4] is subtracted from the sine of the arc formed

[1] [On page 178 Clavius defines prosthaphaeresis as a method in which only addition and subtraction are used.]

[2] [If the arcs are A and B, then the operation will be equivalent to taking sin $(90° - B + A)$.]

[3] [That is, in modern notation

$$1: \sin A = \sin B: \tfrac{1}{2} \sin (90° - B + A),$$

or

$$1: \sin [(90° - B) = \sin B: \tfrac{1}{2} \sin (90° - B + 90° - B),$$

which reduces to

$$1: \cos B = \sin B: \tfrac{1}{2} \sin 2B,$$

or to

$$\sin 2B = 2 \sin B \cos B.]$$

[4] [*I.e.*, $\sin (90° - A - B)$.]

before. Half of the remainder will be the fourth proportional required.[1]

3. If, however, the smaller arc is greater than the complement of the greater (which takes place when the sum of the arcs is greater than the quadrant of a circle), the complement of the greater is subtracted from the lesser arc, so that we again have the difference between the arcs that have been previously added; the sine of this difference is to be added to the sine of the arc previously formed. Half of this sum will be the required fourth proportional.[2]

This is the rule of the above-mentioned author, which will be proven in the following way:

In the first figure we see that EG is the sinus totus. Further, EG is to GK (sine of arc GD), as Ei (sine of arc ID, or HM) to a problematical sine iL. And since the minor arc GD is equal to [itself], DG which is the complement of the greater arc ID (or if GD is the greater and ID the smaller, $ID = DI$, the complement of the greater arc GD), the fourth proportional required will be PQ, which is equal to half the sine MP of the arc MD, composed of the smaller arc DG and of GM, the complement of the greater arc HM.

In the second and third figures we also have the sinus totus EG is to GK (sine of arc GD), as Ei (the sine of the arc IN, or HM) is to the required sine iL. And because in the second figure the smaller arc GD is less than GN, the complement of the greater arc IN (or if by chance GD will be greater and IN smaller, the lesser IN will be smaller than the complement ID of the greater arc GD), the required sine PQ [which is the fourth proportional to I, sin A, sin B] will be obtained in the following way: the sine RP of the difference DN (that is, its equal ME) is to be subtracted from MP, which is the sine of the arc MD composed of the minor arc DG and the complement GM of the greater arc HM. The line

---

[1] [The proportion will then have the form

$$1 : \sin A = \sin B : \tfrac{1}{2}\,[\sin(\overline{90^\circ - A} + B) - \sin(\overline{90^\circ - A} - B)]$$

which is equivalent to

$$1 : \sin A = \sin B : [\tfrac{1}{2} \cos(A - B) - \cos(A + B)].$$

[2] [In modern notation. If $B > 90^\circ - A$ (which means that $A + B > 90^\circ$) we change the way of subtraction, taking $B - (90^\circ - A)$ instead of $(90^\circ - A) - B$.]

PQ will then be half of the remainder *PE*, just as *QR* will be half of the total *MR*.[1]

If, perchance, *GD* is the greater arc and *IN* the smaller, *MP* will nevertheless be the sine of the arc *MB*, composed of the smaller arc *MH* and the complement *HB* of the greater arc *GD*.

In the third figure, since the smaller arc *IN* is greater than *ID*, the complement of the greater arc *GD* (or, if perchance *GD* will be the smaller arc and *IN* the greater, the smaller *GD* will exceed the complement *GN* of the greater arc *IN*), the required fourth proportional will be obtained by adding the sine *RP* of the difference *ND*, that is, adding $ME = RP$ to *MP*, the sine of the arc *MB*, which consists of the smaller arc (*HM*) and *HB*, the complement of the greater. The line *PQ*, which is half of the total line *EP* (since $QR = \frac{1}{2}MR$), will be equal to the required line *iL*.

If the arc *GD* is, perchance, the smaller, and *IN* the greater, *MP* will nevertheless be the sine of arc *MD*, composed of the lesser arc *GD*, and of *GM*, the complement of the greater arc *HM*.

[1] [In modern notation:
$MP = \sin MD = \sin (DG + GM) = \sin (DG + 90^\circ - HM) = \cos (DG - HM)$. Further, triangles $QR = \frac{1}{2}MR$, since in the similar triangles *MIQ* and *MRN* we have $MI = \frac{1}{2}MN$.]

GAUSS

On Conformal Representation

(Translated from the German by Dr. Herbert P. Evans, University of Wisconsin, Madison, Wis.)

Carl Friedrich Gauss was born at Braunschweig April 23, 1777, and died at Göttingen February 23, 1855. From 1795 until 1798 he was a student at Göttingen and during the ten years immediately following this period many of his great fundamental discoveries in pure mathematics and astronomy were made. As a recognition of his work in astronomy he was made director of the Göttingen observatory in 1807, and this position he held until his death. Almost every field of pure and applied mathematics has been enriched by the genius of Gauss, and his researches were so far reaching and fundamental that he is considered as the greatest of German mathematicians. The present memoir was inspired by a prize problem of the Royal Society of Sciences in Copenhagen and is entitled: "General Solution of the Problem to so Represent the Parts of One Given Surface upon another Given Surface that the Representation shall be Similar, in its Smallest Parts, to the Surface Represented." This memoir was written in 1822 and won the prize offered by the Society. It is found in volume 4, pages 192 to 216, of Gauss's collected works, published in 1873.

A transformation whereby one surface is represented upon another with preservation of angles is called today a *conformal*[1] transformation. The earliest conformal transformation dates back to the Greeks, who were familiar with the stereographic projection of a sphere upon a plane. Lagrange[2] considered the conformal representation of surfaces of revolution upon a plane, but it remained for Gauss, in the memoir herein translated, to solve the general problem of the conformal representation of one surface upon another. The memoir may be considered as the basis for the theory of conformal representation and is fundamental to the more modern theory of analytic functions of a complex variable.

Conformal Representation

General Solution of the Problem to so represent the Parts of one Given Surface upon another Given Surface that the representation shall be Similar, in its smallest Parts, to the Surface represented.[3]

---

[1] The term comformal was introduced by Gauss subsquently to the present memoir.

[2] *Collected works* (V. 4, pp. 635–692).

[3] [Gauss, *Werke*, Band 4, p. 193.]

§1. The nature of a curved surface is specified by an equation between the coordinates x, y, z associated with every point on the surface. As a consequence of this equation each of these three variables can be considered as a function of the other two. It is more general to introduce two new variables t, u and to represent each of the variables x, y, z as a function of t and u. By this means definite values of t and u are, at least in general, associated with a definite point of the surface, and conversely.

§2. Let the variables X, Y, Z, T, U have the same significance in reference to a second curved surface as x, y, z, t, u have in reference to the first.

§3. To represent a first surface upon a second is to lay down a law, by which to every point of the first there corresponds a definite point of the second. This will be accomplished by equating T and U to definite functions of the two variables t and u. Insofar as the representation is to satisfy certain conditions these functions can no longer be supposed arbitrary. As thereby X, Y, Z also become functions of t and u, in addition to the conditions which are prescribed by the nature of the two surfaces, these functions must satisfy the conditions which are to be fulfilled in the representation.

§4. The problem of the Royal Society prescribes that the representation should be similar in its smallest parts to the surface represented. First this requirement must be formulated analytically. By differentiation of the functions which express x, y, z, X, Y, Z in terms of t and u there result the following equations:

$$\begin{aligned} dx &= adt + a'du, \\ dy &= bdt + b'du, \\ dz &= cdt + c'du, \\ dX &= Adt + A'du, \\ dY &= Bdt + B'du, \\ dZ &= Cdt + C'du. \end{aligned}$$

The prescribed condition requires, first, that the lengths of all indefinitely short lines extending from a point in the second surface and contained therein shall be proportional to the lengths of the corresponding lines in the first surface, and secondly, that every angle made between these intersecting lines in the first surface shall be equal to the angle between the corresponding lines in the

second surface. A linear element on the first surface may be written

$$\sqrt{(a^2 + b^2 + c^2)dt^2 + 2(aa' + bb' + cc')dt.du + (a'^2 + b'^2 + c'^2)du^2},$$

and the corresponding linear element of the second surface is

$$\sqrt{(A^2+B^2+C^2)dt^2+2(AA'+BB'+CC')dt.du+(A'^2+B'^2+C'^2)du^2}.$$

In order that these two lengths shall be in a given ratio independently of dt and du it is obvious that the three quantities

$$a^2 + b^2 + c^2,\ aa' + bb' + cc',\ a'^2 + b'^2 + c'^2$$

must be respectively proportional to the three quantities

$$A^2 + B^2 + C^2,\ AA' + BB' + CC',\ A'^2 + B'^2 + C'^2.$$

If the endpoints of a second element on the first surface correspond to the values

$$t,\ u \text{ and } t + \delta t,\ u + \delta u,$$

then the cosine of the angle which this element makes with the first element is

$$\frac{(adt + a'du)(a\delta t + a'\delta u) + (bdt + b'du)(b\delta t + b'\delta u) + (cdt + c'du)(c\delta t + c'\delta u)}{\sqrt{\{(adt + a'du)^2 + (bdt + b'du)^2 + (cdt + c'du)^2\}\{(a\delta t + a'\delta u)^2 + (b\delta t + b'\delta u)^2 + (c\delta t + c'\delta u)^2\}}}$$

The cosine of the angle between the corresponding elements on the second surface is given by a similar expression, which is obtained if only a, b, c, a', b', c' are replaced by A, B, C, A', B', C'. Obviously the two expressions become equal if the above mentioned proportionality exists, and the second condition is therefore included in the first, which also is clear in itself by a little reflection.

The analytical expression of our problem is, accordingly, that the equations

$$\frac{A^2 + B^2 + C^2}{a^2 + b^2 + c^2} = \frac{AA' + BB' + CC'}{aa' + bb' + cc'} = \frac{A'^2 + B'^2 + C'^2}{a'^2 + b'^2 + c'^2}$$

shall hold. This ratio will be a finite function of t and u which we will designate by m^2. Then m is the ratio by which linear dimensions on the first surface are increased or diminished in representing them on the second surface (accordingly as m is greater or less than unity). In general this ratio will be different at different points; in the special case for which m is a constant, corresponding finite parts will also be similar, and if moreover $m = 1$ there will be complete equality and the one surface is developable upon the other.

§5. If for the sake of brevity we will put

$$(a^2 + b^2 + c^2)dt^2 + 2(aa' + bb' + cc')dt.du + (a'^2 + b'^2 + c'^2)du^2 = \omega,$$

it is to be noted that the differential equation $\omega = 0$ will allow two integrations. For since the trinomial ω may be broken into two factors linear in dt and du, either one factor or the other must vanish, resulting in two different integrations. One integral will correspond to the equation

$$0 = (a^2 + b^2 + c^2)dt + \{(a'^2 + b'^2 + c'^2 + i\sqrt{(a^2 + b^2 + c^2)(a'^2 + b'^2 + c'^2) - (aa' + bb' + c')^2}\}du$$

(where i for brevity is written in place of $\sqrt{-1}$, since it is readily seen that the irrational part of the expression must be imaginary); the other integral will correspond to a quite similar equation, obtained by exchanging $-i$ with i. Consequently if the integral of the first equation is

$$p + iq = \text{Const.},$$

where p and q signify real functions of t and u, the other integral will be

$$p - iq = \text{Const.}$$

Consequently, by the nature of the case,

$$(dp + idq)(dp - idq)$$

or

$$dp^2 + dq^2$$

must be a factor of ω, or

$$\omega = n(dp^2 + dq^2),$$

where n will be a finite function of t and u.

We will now designate by Ω the trinomial into which

$$dX^2 + dY^2 + dZ^2$$

changes when for dX, dY, dZ are substituted their values in terms of T, U, dT, dU, and suppose that as in the foregoing the two integrals of the equation $\Omega = 0$ are

$$P + iQ = \text{Const.},$$
$$P - iQ = \text{Const.},$$

and

$$\Omega = N(dP^2 + dQ^2),$$

where P, Q, N are real functions of T and U. These integrations (aside from the general difficulties of integration) can obviously be carried out previous to the solution of our main problem.

If now such functions of t, u be substituted for T, U that the condition of our main problem is satisfied, then Ω may be replaced by $m^2\omega$ and we have

$$\frac{(dP + idQ)(dP - idQ)}{(dp + idq)(dp - idq)} = \frac{m^2 n}{N}.$$

It is easily seen however, that the numerator in the left hand side of this equation can be divisible by the denominator only if either

$$dP + idQ \text{ is divisible by } dp + idq$$

and

$$dP - idQ \text{ is divisible by } dp - idq$$

or

$$dP + idQ \text{ is divisible by } dp - idq$$

and

$$dP - idQ \text{ is divisible by } dp + idq.$$

In the first case therefore, $dP + idQ$ will vanish if $dp + idq = 0$, or $P + iQ$ will be constant if $p + iq$ is constant, *i. e.*, $P + iQ$ will be a function only of $p + iq$; and likewise $P - iQ$ will be a function only of $p - iq$. In the other case $P + iQ$ will be a function of $p - iq$ only and $P - iQ$ a function of $p + iq$. It is easily understood that these conclusions also hold conversely; namely that, if $P + iQ$, $P - iQ$ are assumed to be functions of $p + iq$, $p - iq$ (either respectively or reversely), the finite divisibility of Ω by ω follows and accordingly the required proportionality exists.

Moreover, it is easily seen that if, for example, we put

$$P + iQ = f(p + iq),$$

and

$$P - iQ = f'(p - iq),$$

the nature of the function f' is dependent upon that of f. That is, if the constant quantities which the latter may perhaps involve are all real, then f' must be identical with f, in order for real values of P, Q to correspond to real values of p, q. On the contrary supposition the function f' may be obtained from f by merely substituting $-i$ for i therein. Accordingly we have

$$P = \tfrac{1}{2}f(p + iq) + \tfrac{1}{2}f'(p - iq),$$
$$iQ = \tfrac{1}{2}f(p + iq) - \tfrac{1}{2}f'(p - iq),$$

or, what is the same thing, when the function f is assumed quite arbitrary (constant imaginary elements included at pleasure), P is placed equal to the real part and iQ (in the case of the second

solution $-iQ$) to the imaginary part of $f(p + iq)$, then by eliminating T and U they will be expressed as functions of t and u. Thus the given problem is completely and generally solved.

§6. If $p' + iq'$ represents an arbitrary function of $p + iq$ (where p', q' are real functions of p, q), it is easily seen that also

$$p' + iq' = \text{Const. and } p' - iq' = \text{Const.}$$

represent integrals of the differential equation $\omega = 0$; in fact, these equations are quite the equivalents of

$$p + iq = \text{Const. and } p - iq = \text{Const.}$$

respectively. Similarly the integrals

$$P' + iQ' = \text{Const. and } P' - iQ' = \text{Const.}$$

of the differential equation $\Omega = 0$ will be the equivalents of

$$P + iQ = \text{Const. and } P - iQ = \text{Const.}$$

respectively, if $P' + iQ'$ represents an arbitrary function of $P + iQ$ (where P', Q' are real functions of P and Q). From this it is clear that in the general solution of our problem, which has been given in the foregoing section, $p'q'$ can take the place of p, q and P', Q' the place of P, Q, respectively. Although the solution gains no greater generality by this substitution, yet occasionally in the application, one form can be more useful than the other.

§7. If the functions which arise from the differentiation of the arbitrary functions f, f' are designated by φ and φ' respectively, so that

$$df(v) = \varphi(v)dv, \qquad df'(v) = \varphi'(v)dv,$$

then as a result of our general solution it follows that

$$\frac{dP + idQ}{dp + idq} = \varphi(p + iq), \qquad \frac{dP - idQ}{dp - idq} = \varphi'(p - iq).$$

Therefore

$$\frac{m^2 n}{N} = \varphi(p + iq) . \varphi'(p - iq).$$

The ratio of magnification is consequently defined by the formula

$$m = \sqrt{\frac{dp^2 + dq^2}{\omega} \cdot \frac{\Omega}{dP^2 + dQ^2} \cdot \varphi(p + iq) \cdot \varphi'(p - iq)}.$$

§8. We will now illustrate our general solution by means of several examples,[1] whereby the kind of application, as well as the

[1] [Only the first example is reproduced here, sections 9–13 of the original memoir being omitted.]

nature of several details still to come in for consideration, will best be brought to light.

First consider two plane surfaces, in which case we may write

$$x = t,\ y = u,\ z = 0,$$
$$X = T,\ Y = U,\ Z = 0.$$

The differential equation

$$\omega = dt^2 + du^2 = 0$$

gives here the two integrals

$$t + iu = \text{Const.}, \qquad t - iu = \text{Const.},$$

and likewise the two integrals of the equation

$$\Omega = dT^2 + dQ^2 = 0$$

are as follows:

$$T + iU = \text{Const.}, \qquad T - iU = \text{Const.}$$

The two general solutions of the problem are accordingly

$$\text{I. } T + iU = f(t + iu), \qquad T - iU = f'(t - iu).$$
$$\text{II. } T + iU = f(t - iu), \qquad T - iU = f'(t + iu).$$

These results may also be expressed thus: If the characteristic f designates an arbitrary function, the real part of $f(x + iy)$ is to be taken for X, and the imaginary part, with omission of the factor i, for either Y or $-Y$.

If the notation φ, φ' are used in the sense of §7, and if we put

$$\varphi(x + iy) = \xi + i\eta, \qquad \varphi'(x - iy) = \xi - i\eta,$$

where obviously ξ and η are to be real functions of x and y, then in the case of the first solution we have

$$dX + idY = (\xi + i\eta)(dx + idy),$$
$$dX - idY = (\xi - i\eta)(dx - idy),$$

and consequently

$$dX = \xi dx - \eta dy,$$
$$dY = \eta dx + \xi dy.$$

Now take

$$\xi = \sigma \cos \gamma, \qquad \eta = \sigma \sin \gamma$$
$$dx = ds \,.\, \cos g, \qquad dy = ds \,.\, \sin g$$
$$dX = dS \,.\, \cos G, \qquad dY = dS \,.\, \sin G.$$

thereby defining ds as a linear element in the first plane making an angle g with the x-axis and dS as the corresponding linear

element in the second plane making an angle G with the X-axis. From these equations there results

$$dS \,.\, \cos G = \sigma ds \cos (g + \gamma),$$
$$dS \,.\, \sin G = \sigma ds \sin (g + \gamma),$$

and if σ is regarded as positive (which is permissible) it follows that

$$dS = \sigma ds,\ G = g + \gamma$$

It is thus seen (in agreement with §7) that σ represents the ratio of magnification of the element ds in the representation dS, and as requisite, is independent of g; also, since γ is independent[1] of g, it follows that all linear elements extending from a point in the first plane are represented by elements in the second plane, which meet at the same angles, *in the same sense*, as do the corresponding elements in the first plane.

If f is taken to be a linear function, so that $f(v) = A + Bv$, where the constant coefficients are of the form

$$A = a + bi,\ B = c + ei$$

then

$$\varphi(v) = B = c + ei$$

and consequently[2]

$$\sigma = \sqrt{c^2 + e^2},\ \gamma = \text{arc tan}\,\frac{e}{c}.$$

The ratio of magnification is therefore the same at all points and the representation is completely similar to the surface represented.[3] For every other function f (as one can easily prove) the ratio of magnification will not be constant, and the similarity will therefore occur only in the smallest parts.

If points in the second plane are prescribed which, in the representation, are to correspond with a certain number of given points in the first plane, then by the common method of interpolation we can easily find the simplest algebraic function f for which this condition is satisfied. Namely, if we designate the values of

---

[1] [That γ is independent of g follows from the fact that σ and ξ, η are independent of g.]

[2] [This follows from the fact that, in this case, $\xi = c$, $\eta = e$, and by definition $\sigma = \sqrt{\xi^2 + \eta^2}$.]

[3] [The similarity is called complete if finite parts of the two surfaces are similar.]

$x + iy$ for the given points by a, b, c, etc., and the corresponding values of $X + iY$ by A, B, C etc., then we have to make

$$f(v) = \frac{(v-b)(v-c)\ldots}{(a-b)(a-c)\ldots} \,.\, A + \frac{(v-a)(v-c)\ldots}{(b-a)(b-c)\ldots} \,.\, B + \frac{(v-a)(v-b)+\ldots}{(c-a)(c-b)+\ldots} \,.\, C + \text{etc.},$$

which is an algebraic function of v whose order is a unit less than the number of given points.[1] In the case of only two points, the function becomes linear and consequently there is complete similarity.

If the second solution is carried through in the same way we find that the similarity is reversed, as all elements in the representation make the same angles with one another as do the corresponding elements in the original surface but in the reverse sense, and so that lies to the right which before lay to the left. This difference is not essential however, and disappears if we take for the under side in one plane the side before regarded as the upper side.

. . . .

§14. It remains to consider more fully one feature occurring in the general solution. We have shown in §5, that there are always just two solutions, since either $P + iQ$ must be a function of $p + iq$, and $P - iQ$ a function of $p - iq$; or $P + iQ$ must be a function of $p - iq$, and $P - iQ$ a function of $p + iq$. We shall now show that always in the case of one solution the parts in the representation are situated similarly as on the surface represented; in the other solution, on the contrary, they lie in the reverse sense; at the same time we shall specify the criterion by means of which this can be settled a priori.

First of all we observe, that of perfect or reversed similarity there can be discussion only insofar as on each of the two surfaces two sides are distinguished, one of which is considered as the upper and the other as the under. Since this in itself is somewhat arbitrary, the two solutions do not differ at all essentially, and a reversed similarity becomes perfect as soon as the side on one surface, considered as the upper side, is taken as the under side. In our solution this distinction cannot present itself, since the surfaces were defined only by the coordinates of their points. If

[1] [In the original memoir an application of this process to geodesy is also mentioned.]

one is concerned with this distinction, the nature of the surfaces must first be specified in another manner which includes this in itself. For this purpose we shall assume that the nature of the first surface is defined by the equation $\psi = 0$, where ψ is a given one-valued function of x, y, z. At all points of the surface the value of ψ will thus be zero, and at all points not on the surface it will have a value different than zero. By a passage through the surface, generally speaking, ψ will change from positive to negative or by opposite motion from negative to positive, *i. e.*, on one side the value of ψ will be positive and on the other negative: the first side will be considered as the upper and the other as the under. Likewise the nature of the second surface is to be similarly specified by the equation $\Psi = 0$, where Ψ is a given one-valued function of the coordinates X, Y, Z. Differentiation gives

$$d\psi = e\,dx + g\,dy + h\,dz,$$
$$d\Psi = E\,dx + G\,dY + H\,dZ,$$

where e, g, h will be functions of x, y, z and E, G, H functions of X, Y, Z.

Since the considerations through which we attain our aim, although in themselves not difficult, are yet of a somewhat unusual kind, we shall take pains to give them the greatest clarity. We shall assume six intermediate representations in the plane to be inserted between the two corresponding representations on the surfaces whose equations are $\psi = 0$ and $\Psi = 0$, so that eight different representations come in for consideration, namely the surfaces:

	The corresponding points of which have as coordinates:
1. The original in the surface $\psi = 0$	x, y, z.
2. Representation in the plane	$x, y, 0$.
3. Representation in the plane	$t, u, 0$.
4. Representation in the plane	$p, q, 0$.
5. Representation in the plane	$P, Q, 0$.
6. Representation in the plane	$T, U, 0$.
7. Representation in the plane	$X, Y, 0$.
8. Representation in the surface $\Psi = 0$	X, Y, Z.

We shall now compare these different representations solely in relation to the relative positions of their infinitesimal linear elements, disregarding entirely the ratio of their lengths; two representations will be considered as similar, if two linear elements extending from a point are such that the one which lies to the right

in the one surface corresponds to the one which lies to the right in the other; in the contrary case the linear elements will be said to be reversely situated. In case of the planes 2–7 the side on which the third coordinate has a positive value will always be considered as the upper side; in case of the first and last surfaces, on the other hand, the distinction between the upper and under sides depends merely upon the positive or negative values of ψ and Ψ, as has already been agreed upon.

First of all, it is clear that at every point of the first surface where one arrives at the upper side by giving z a positive increment, for x and y unchanged, the representation in 2 will be similar to that in 1; this will obviously be the case whenever b is positive; and the contrary will occur when b is negative, in which case the representation 2 will be reversely situated with respect to 1.

In the same way the representations 7 and 8 will be situated similarly or reversely, accordingly as H is positive or negative.

In order to compare the representations in 2 and 3 let ds be the length of an infinitesimal line in the former surface, extending from the point with coordinates x, y to another with coordinates $x + dx$, $y + dy$, and let l denote the angle between this element and the positive x-axis, the angle increasing in the same sense as we pass from the x-axis to the y-axis; thus:

$$dx = ds \,.\, \cos l, \qquad dy = ds \,.\, \sin l.$$

In the representation in 3 let $d\sigma$ be the length of the line which corresponds to ds and let λ, in the above sense, be the angle it makes with the t-axis, so that

$$dt = d\sigma \,.\, \cos \lambda, \qquad du = d\sigma \,.\, \sin \lambda.$$

We have therefore, in the notation of §4,

$$ds \,.\, \cos l = d\sigma \,.\, (a \cos \lambda + a' \sin \lambda),$$
$$ds \,.\, \sin l = d\sigma \,.\, (b \cos \lambda + b' \sin \lambda),$$

and consequently

$$\tan l = \frac{b \cos \lambda + b' \sin \lambda}{a \cos \lambda + a' \sin \lambda}.$$

If x and y are now considered as fixed and l, λ as variable, it follows by differentiation that

$$\frac{dl}{d\lambda} = \frac{ab' - ba'}{(a \cos \lambda + a' \sin \lambda)^2 + (b \cos \lambda + b' \sin \lambda)^2} = (ab' - ba')\left(\frac{d\sigma}{ds}\right)^2.$$

It is thus seen that, accordingly as $ab' - ba'$ is positive or negative, l and λ will increase simultaneously or change in the opposite sense, and therefore in the first case the representations 2 and 3 are similarly situated, while in the second case they are reversely situated.

From the combination of these results with the foregoing it follows that the representations 1 and 3 are similarly or reversely situated, accordingly as $(ab' - ba')/h$ is positive or negative.

Since the equation

$$edx + gdy + hdz = 0$$

as also

$$(ea + gb + hc)dt + (ea' + gb' + hc')du = 0,$$

must hold on the surface $\psi = 0$, irrespective of how the ratio of dt and du is chosen, we have identically

$$ea + gb + hc = 0 \text{ and } ea' + gb' + hc' = 0.$$

Wherefrom it follows that e, g, h must be respectively proportional to the quantities $bc' - cb'$, $ca' - ac'$, $ab' - ba'$, thus

$$\frac{bc' - cb'}{e} = \frac{ca' - ac'}{g} = \frac{ab' - ba'}{h}.$$

We can apply any one of these three expressions, or, on multiplication by the positive quantity $e^2 + g^2 + h^2$, the resulting symmetrical expression,

$$ebc' + gca' + hab' - ecb' - gac' - hba',$$

as a criterion for the similarity or reversal of position of the parts in the representations 1 and 3.

Likewise the similarity or reversal of parts in the representations 6 and 8 depends upon the positive or negative value of the quantity

$$\frac{BC' - CB'}{E} = \frac{CA' - AC'}{G} = \frac{AB' - BA'}{H},$$

or, if we prefer, upon the sign of the symmetrical quantity

$$EBC' + GCA' + HAB' - ECB' - GAC' - HBA'.$$

The comparison of the representations 3 and 4 is based on quite similar grounds as that of 2 and 3, and the similar or reverse situation of the parts depends upon the positive or negative sign of the quantity

$$\frac{\partial p}{\partial t}\cdot\frac{\partial q}{\partial u} - \frac{\partial p}{\partial u}\cdot\frac{\partial q}{\partial t}.$$

Likewise the positive or negative sign of

$$\frac{\partial P}{\partial T}\cdot\frac{\partial Q}{\partial u}-\frac{\partial P}{\partial U}\cdot\frac{\partial Q}{\partial T}$$

determines the similar or reverse situation of the parts in the representations 5 and 6.

Finally, to compare the representations 4 and 5 the analysis of §8 may be employed, from which it is clear that these are similar or reversed in the situation of their smallest parts, accordingly as the first or second solution is chosen, that is, whether

$$P+iQ=f(p+iq) \text{ and } P-iQ=f'(p-iq),$$

or

$$P+iQ=f(p-iq) \text{ and } P-iQ=f'(p+iq).$$

From all this we now conclude that if the representation in the surface $\Psi=0$ is not only to be similar in its smallest parts to its image on the surface $\psi=0$, but similar in position as well, attention must be paid to the number of the four quantities

$$\frac{ab'-ba'}{b},\quad \frac{\partial p}{\partial t}\cdot\frac{\partial q}{\partial u}-\frac{\partial p}{\partial u}\cdot\frac{\partial q}{\partial t},\quad \frac{\partial P}{\partial T}\cdot\frac{\partial Q}{\partial U}-\frac{\partial P}{\partial U}\frac{\partial Q}{\partial T},\quad \frac{AB'-BA'}{H},$$

which have negative signs. If none or an even number of them have negative signs the first solution must be chosen; if one or three of them have negative signs, the second solution must be chosen. For any other choice the similarity is always reversed.

Moreover it can be shown that, if the above four quantities are designated by r, s, S, R respectively, the equations

$$\frac{r\sqrt{e^2+g^2+b^2}}{s}=\pm n,\ \frac{R\sqrt{E^2+G^2+H^2}}{S}=\pm N$$

always hold, where n and N have the same significance as in §5; we omit the easily found proof of this theorem here, however, since this, for our purpose, is not necessary.

STEINER

On Birational Transformations between Two Spaces

(Translated from the German by Professor Arnold Emch, University of Illinois, Urbana, Ill.)

Jakob Steiner (1796–1863) was born in humble circumstances and could not write before he reached the age of fourteen. Pestalozzi (1746–1827) took him into his school at Yverdon, Switzerland, at the age of seventeen and inspired in him a love for mathematics. He went to the University of Heidelberg in 1818 and in 1834 became a professor in the University of Berlin. He was a prolific writer on geometry. In his classic *Systematische Entwickelung der Abhängikeit geometrischer Gestalten von einander* (Berlin, 1832) he established and discussed (pp. 251–270) the so-called skew projection (Schiefe Projektion) and its applications. This projection is based upon two fixed planes, (x) and (x'), and two fixed axes, l and y in space. From every point x in (x) there is, in general, one transversal through l and y which cuts (x') in a point x'. Thus to every point in (x) there corresponds a point in (x'), and conversely. To lines correspond conics, etc. By this construction there is established a general quadratic transformation between two planes, with distinct real fundamental points and lines in both planes. On page 295, Steiner indicates the quadratic transformation between two spaces, and in a footnote he makes the significant statement quoted below, thus clearly realizing the possibility of transformations of higher order, including Cremona transformations beyond the quadratic. For a further discussion see "Selected Topics in Algebraic Geometry," *Bulletin of the National Research Council* (Washington, 1928, Chap. I).

How in this manner other more complex systems of this kind may be established is easily seen. Namely, by every porism in which, for example, the relation between two points is such that, while one of the points describes a line (or a curve), the other describes a definite curve, such a system arises...

CREMONA

On the Geometric Transformations of Plane Figures

(Translated from the Italian by E. Amelotti, M.S. University of Illinois, Urbana, Ill.)

Luigi Cremona was born in Pavia Dec. 7, 1830, and died in Rome June 10, 1903. In 1860 he became professor of higher geometry in Bologna, in 1866 professor of geometry and graphical statistics at Milan, and in 1873 professor of higher mathematics and director of engineering schools in Rome.

Synthetic geometry was studied by him with great success. A memoir in 1866 on cubic surfaces secured half of the Steiner prize from Berlin. He wrote on plane curves, on surfaces, and on birational transformations of plane and solid space. A Cremona transformation is equivalent to a succession of quadratic transformations of Magnus's type. Cremona's theory of transformation of curves was extended by him to three dimensions. For further information concerning the life and works of Cremona the reader is referred to the *Periodico di Matematica per l'Insegnamento Secondario*, Ser. 1, Vol. 5–6, 1890–1891; *Supplemento*, 1901–1902, pp. 113–114; and the *History of Mathematics*, by Florian Cajori, New York, 1926 ed.

In modern algebraic geometry such properties of figures are studied as are invariant under (*a*) the projective transformation, (*b*) the Cremona transformation, or (*c*) the birational transformation. The first clear survey of the aggregate of Cremona transformations in the plane is contained in the article here reported. In this and in later memoirs of Cremona the fundamental properties of such transformations in plane and space are established.

The article here translated is taken from the *Giornale di Matematiche*, of Battaglini, vol. I, ser. 1 (1863), pp. 305–311.

Messrs. Magnus and Schiaparelli, the one in Tome 8 of Crelle's *Journal*, the other in a very recent volume of the memoirs of the Accademia Scientifica di Torino, were seeking for the analytic expression for the geometric transformation of a plane figure into another plane figure under the condition that to any point of one there corresponds only one point of the other, and conversely to each point of the other only one point of the first (*transformation of the first order*). And from the above cited authors it seems that one should conclude that, in the most general situation to the lines of one figure there corresponds, in the other, conics circumscribed about a fixed triangle (real or imaginary), *i. e.*, that the most general transformation of the first order is that which Schiaparelli calls *conical transformation.*

But it is evident that by applying to a figure a succession of conical transformations there will result from this composition a transformation which is still of the first order, even though in it, to the right lines of the given figure there would correspond in the transformed plane not conics, but curves of higher order.

. . . .

Upon the Transformation of Plane Figures

I will consider two figures, one located in a plane P, the other in a plane P', and will suppose that the second was deduced from the first by means of any law of transformation although in such manner that to each point of the first figure there correspond only one of the second, and conversely.

The geometric transformations subject to the conditions above mentioned are the only ones which I will examine in this account: and they shall be called "transformations of the first order"[1] to distinguish them from others which are determined by different conditions.

Assuming that the transformation by means of which the proposed figures are deduced, one from the other, are among those of the first order the most general, I then ask the question: What curve of a figure corresponds to right lines of the other?

Let n be the order of the curve which in the plane P' (or P) corresponds to any line whatsoever of the plane P (or P'). Since a line of the plane P is determined by two points a, b, then the two corresponding points a', b', of the plane P' are sufficient to determine the curve which corresponds to the given line. Therefore the curves of the one figure corresponding to the lines of the other form a system such that through two arbitrarily given points only one line passes through them; *i.e.*, those curves form a geometric net of order n (II).

A curve of order n is determined by $\frac{1}{2}n(n+3)$ conditions; therefore the curve of a figure corresponding to right lines of another are subjected to $\frac{1}{2}n(n+3) - 2 = \frac{1}{2}(n-1)(n+4)$ common conditions.

Two right lines of the one figure have only one point in common, a, determined by them. The point a' corresponding to a will

---

[1] Schiaparelli: "Sulla Trasformazione Geometrica delle Figure ed in Particolare sulla Trasformazione iperbolica" (*Memorie della R. Accademia delle Scienze di Torino*, serie, 2^a, tomo XXI, Torino 1862).

belong to the two curves of order n to which the two lines correspond. And since these two curves must determine the point a', the remaining $n^2 - 1$ intersections must be common to all the curves of the geometric net above mentioned.

Let x_r be the number of r-ple (multiple points of order r) points common to these curves; since an r-ple point common to two curves is equivalent to r^2 intersections of the same, then we will have evidently:

(1) $$x_1 + 4x_2 + 9x_3 + \ldots + (n-1)^2 x_{n-1} = n^2 - 1.$$

The $x_1 + x_2 + x_3 + \ldots + x_{n-1}$ points common to the curves of the net constitute the $\frac{1}{2}(n-1)(n+4)$ conditions which determine it. If a curve must pass r times through a given point, that is equivalent to $\frac{1}{2}r(r+1)$ conditions;

(2) $$x_1 + 3x_2 + 6x_3 + \ldots + \tfrac{1}{2}n(n-1)x_{n-1} = \tfrac{1}{2}(n-1)(n+4).$$

Equations (1) and (2) are evidently the only conditions which the integral positive numbers $x_1, x_2, \ldots, x_{n-1}$ must satisfy[1] (or P) corresponds to any line whatsoever of the plane P (or P'). Since a line of the plane P is determined by two points a, b, then the two corresponding points a', b' of the plane P' are sufficient to determine the curve which corresponds to the given line. Therefore the curves of the one figure corresponding to the lines of the other form a system such that through two arbitrarily given points only one line passes through them; *i. e.*, those curves form a geometric net of order n.[2]

A curve of order n is determined by $\frac{1}{2}n(n+3)$ conditions; therefore the curve of a figure corresponding to right lines of another are subjected to $\frac{1}{2}n(n+3) - 2 = \frac{1}{2}(n-1)(n+4)$ common conditions.

Two right lines of the one figure have only one point in common a, determined by them. The point a' corresponding to a will belong to the two curves of order n to which the two lines correspond. And since these two curves must determine the point a', the remaining $n^2 - 1$ intersections must be common to all the curves of the geometric net above mentioned.

---

[1] [Cremona then inserts a footnote explaining how it is that one does not obtain new equations when one considers the curves which in the plane P' correspond to curves of a given order μ in the plane P.]

[2] See my "Introduzione ad una teoria geometrica delle curve piane, Page 71."

Let x_r be the number of r-ple (multiple points of order r) points common to these curves; since an r-ple point common to two curves is equivalent to r^2 intersections of the same then we will have evidently:

$$(1) \qquad x_1 + 4x_2 + 9x_3 + \ldots + (n-1)^2 x_{n-1} = n^2 - 1$$

The $x_1 + x_2 + x_3 + \ldots + x_{n-1}$ points common to the curves of the net constitute the $\frac{(n-1)(n+4)}{2}$ conditions which determine it. If a curve must pass r times through a given point, that is equivalent to $\frac{r(r+1)}{2}$ conditions; therefore

Examples.—For $n = 2$, the equations (1) and (2) reduce to the single equation $x_1 = 3$; *i. e.*, to the lines of a figure there will correspond in the other curves of second order circumscribed about a fixed triangle.

This is the aforesaid "Conical Transformation" considered by Steiner,[1] by Magnus,[2] and by Schiaparelli.[3]

For $n = 3$, one has, from (1) and (2),

$$x_1 = 4,\ x_2 = 1;$$

i. e., to the right lines of the one figure there correspond in the other curves of third order all having a double and four simple points in common.

For $n = 4$, (1) and (2) become

$$x_1 + 4x_2 + 9x_3 = 15,$$
$$x_1 + 3x_2 + 6x_3 = 12,$$

which admit the two solutions

$$\textit{First:}\quad x_1 = 3,\ x_2 = 3,\ x_3 = 0;$$
$$\textit{Second:}\ x_1 = 6,\ x_2 = 0,\ x_3 = 1;$$

etc.

On eliminating x_1 from the equations (1) and (2) one obtains the following:

$$(3) \quad x_2 + 3x_3 + \ldots + \frac{(n-1)(n-2)}{2}x_{n-1} = \frac{(n-1)(n-2)}{2}$$

from which one sees that x_{n-1} can not have other than one of these two values:

$$x_{n-1} = 1,\ x_{n-1} = 0$$

[1] *Systematische Entwickelung*, u.s.w., Berlin, 1832, page 251. [See page 476.]

[2] *Crelle's Journal*, t. 8, page 51.

[3] *Loco citato.*

and that in the case $x_{n-1} = 1$ one necessarily has:

$$x_2 = 0, x_3 = 0, \ldots, x_{n-2} = 0$$

and by virtue of (1) $x_1 = 2(n - 1)$

I propose to prove that the transformation corresponding to these values of $x_1, x_2, \ldots, x_{n-1}$ is, for an arbitrary value of n, geometrically possible.

Let it be supposed that the two figures be located in two distinct planes P, P', in such way that to each point of the first plane there corresponds a unique point of the second, and conversely. I will imagine two directrix curves such that through an arbitrary point of space it will be possible to pass only one line to meet both, and I will consider as correspondents the points in which this line meets the planes P, P'.

Let p and q be the orders of the two directrix curves and r the number of their common points. Assuming an arbitrary point 0 of the space as the vertex of two cones, the directrices of which are the above given curves the orders of these two cones will be p, q and therefore they will have $p.q$ common generators. Included among these are the lines which unite O with the r points common to the two directrix curves, and the remaining $pq - r$ generators common to the two cones will be, consequently, the right lines that from O can be drawn to meet both the one and the other directrix curve. But the lines endowed with such property we wish reduced to only one; therefore it must be true that

$$pq - r = 1 \tag{4}$$

Furthermore to any line R situated in one of the planes P, P', there will correspond in the other a curve of order n; *i. e.*, a variable line which meets constantly the line R and the two directrix curves of order p, q must generate a warped surface of order n. One seeks therefore the order of the surface generated by a variable line which cuts three given directrices, the first of which is a line R, and the other two of order p, q have r points in common. The number of the lines which meet three given lines and a curve of order p is $2p$: this being the number of points common to the curve of order p and to the hyperboloid which has as directrices the three given lines. This amounts to saying that $2p$ is the order of a warped surface the directrices of which are two curves and a curve of order p. This surface is met by the curve of order q in $2.p.q - r$ points not situated on the curve of order p.

Therefore the order of the warped surface which has for directrices a line and curves of order p, q, having r common points, is $2pq - r$. Therefore we must have:

$$(5) \qquad 2pq - r = n$$

From the equations (4) and (5) one gets

$$(6) \qquad p.q = n - 1, \quad r = n - 2$$

Let it be supposed that the line R is in plane P, and consider the corresponding curve of order n in plane P', *i. e.*, the intersection of this plane with the warped surface of order $2p.q - r$ previously mentioned. The curve of which one is dealing will have:

p multiple points of order q; they are the intersections of the plane P' with the directrix curve of order p (in fact from each point of this curve it is possible to draw q lines to meet the other directrix curves and the line R, or in other words the directrix curve of order p is multiple of order q on the warped surface);

q multiple points of order p, and they are the intersections of the plane P' with the directrix curve of order q (because analogously this one is multiple of order p on the warped surface);

$p.q$ simple points of intersection of the right line common to the planes P', P, with the lines which from the points where the directrix of order p cuts the plane P, go to the points where the other directrix cuts the same plane.

These $p + q + pq$ points do not vary, as R varies, *i. e.*, they are points common to all the curves of order n, of plane P', corresponding to the lines of plane P. Therefore we will have:

$$x_1 = p.q, \; x_p = q, \; x_q = p.$$

and the other x's will be equal to zero; thus the equations (1) and (2) give, having regard to the first of (6):

$$p + q = n$$

And this one combined with the first of (6) gives as a result

$$p = n - 1, \; q = 1$$

This signifies that of the two directrices, one will be a curve of order $n - 1$ and the other a line which will have $n - 2$ points in common. This condition can be verified by a line and a plane curve of order $n - 1$ (not situated in the same plane) provided that the latter have a multiple point of order $n - 2$, and the directrix passes through this multiple point.

Also, the directrix of order $n-1$ can be a twisted curve; because, for example, on the surface of an hyperboloid one can describe[1] a twisted curve K of order $n-1$ which will be met by each of the generators of same system in $n-2$ points (and in consequence by each generator of the other system in only one point). We can therefore assume such twisted curve and a generator D of the first system as directrices of the transformation.

In this transformation, to each point a of the plane P there corresponds one and only one point a' of the plane P', and conversely, which point a' one determines thus. The plane drawn through the point a and through the line D meets the curve K in only one point outside of the line D. This point joined to a gives a line which meets the plane P' in the required point a'.

If R is any line in plane P, the warped surface (of order n) which has as directrices the lines K, D, R, cuts the plane P' in the curve (of order n) corresponding to R. All the curves which analogously correspond to lines have in common a multiple point of order $n-1$ and $2(n-1)$ simple points, *i. e.*, First, the point in which D meets the plane P'; Second, the $n-1$ points in which the plane P' is met by the directrix K; Third, the $n-1$ points in which the line of intersection of P, P' is met by the lines which unite the point common to the line D and the plane P with points common to the curve K and the same plane P.

In other words: The warped surface analogous to that one the directrices of which are K, D, R, all have in common: First, The directrix D (multiple of order $n-1$, and thus equivalent to $(n-1)^2$ common lines); Second, The curvilinear (simple) directrix K; Third, $n-1$ generators (simple) situated in the plane P. All these curves taken together are equivalent to a curve of order $(n+1)^2+2(n-1)$. Therefore two warped surfaces (of order n) determined by two lines R, S, in the plane P, will have also in common a line; which evidently unites the point a, of intersection of R, S with the corresponding point a', common to the two curves which in the plane P' correspond to the lines R, S.

If the line R goes through the point d in which point D meets the plane P it is evident that the relative ruled surface decomposes into a cone which has its vertex at d and as directrix the curve K, and into the plane which contains the lines D, R.

If the line passes through one of the points k common to the plane P and curve K, the relative ruled surface decomposes into

[1] Comptes rendus de l' Académie de France, 24 Juin, 1861.

the plane which contains the point k and the line D, and into the warped surface of order $n - 2$, having directrices K, D, R.

If the line R passes through two of the points k, the relative ruled surface will decompose into two planes and into a warped surface of order $n - 2$.

And it is also very easy to see that any curve C, of order μ, given in the plane P, gives rise to a warped surface of order μn, for which D is multiple of order $\mu(n - 1)$ and K is multiple of order μ. Therefore to the curve C there will correspond in the plane P' a curve of order μn, having: First; A multiple point of order $\mu(n - 1)$ upon D; Second; $n - 1$ multiple points of order μ, upon K; Third; $n - 1$ multiple points of order μ, upon the line which is a common intersection of P, P'.

Applying to the aforesaid things the principle of duality we will obtain two figures: one composed of lines and planes passing through the point 0; the other of lines and planes passing through another point 0′; and the two figures will have such relations to each other, that to each plane of one there will correspond only one plane of the other and conversely; and to the lines of any one of the figures there will correspond in the other a conical surface of class n, having in common $x_1, x_2, \ldots, x_{n-1}$ tangent planes simple and multiple. The numbers $x_1, x_2, \ldots, x_{n-1}$ will be connected by the same equations (1) and (2).

In particular then, to deduce one figure from the other we can assume as directrices a fixed line D and a developable surface K of class $n - 1$, which has $n - 2$ tangent planes passing through D. Then, given any plane π through 0 which cuts D in a point a; through this point there passes (other than the $n - 2$ planes through D) only one tangent plane which will cut π along a certain line. The plane π' determined by it and the point 0′ is the correspondent of π.

Cutting then the two figures with two planes P and P' respectively, we will obtain in these, two figures such that to each right line of one there will correspond a single right line in the other and conversely; but to a point of the one of the two planes there will correspond in the other a curve of class n, having a certain number of fixed, simple and multiple tangent lines.

LIE

On a Class of Geometric Transformations

(Translated from the Norwegian by Professor Martin A. Nordgaard, St. Olaf College, Northfield, Minn.)

Marius Sophus Lie (Dec. 17, 1842–Feb. 18, 1899) was the most prominent Scandinavian mathematician of his time. He lived for a time in France, but at the age of thirty became professor of mathematics at Christiania (Oslo) and from 1886 to 1898 he held a similar position at Leipzig. Of his style of discourse Klein has this to say:

"To fully understand the mathematical genius of Sophus Lie, one must not turn to the books recently published by him in collaboration with Dr. Engel, but to his earlier memoirs, written during the first years of his scientific career. There Lie shows himself the true geometer that he is, while in his later publications, finding that he was but imperfectly understood by the mathematicians accustomed to the analytical point of view, he adopted a very general analytical form of treatment that is not always easy to follow."[1]

Lie's earliest writings, when his ideas, as Klein says, were still in their "nascent" stage, possess a vividness and a happy directness of expression that is not always noticeable in his later exposition.

It was in 1869–1870, while still a young man, that he made the remarkable discovery of a contact transformation by which a sphere can be made to correspond to a right line. He communicated the results of his discovery to the Christiania Academy of Sciences in July and October, 1870, in a memoir entitled "Over en Classe geometriske Transformationer." The memoir is published in the society's *Proceedings* for 1871, pp. 67–109; the translation of which is here presented. It is because of a general impression that the German version was lacking in the force of the original that it was decided to present the memoir through a direct translation from the Norwegian instead of relying upon the one later published in Berlin.

INTRODUCTION[2]

The rapid development of geometry in the present century has been closely related to and dependent on the philosophic views

[1] Felix Klein, in his lecture on Mathematics, at the Evanston Colloquium, 1893. Macmillan, New York.

[2] The most important points of view in this memoir were communicated to the Christiania Academy of Sciences in July and October, 1870. Compare a note by Mr. Klein and myself in the Berlin Academy's *Monatsbericht* for Dec. 15, 1870.

of the nature of Cartesian geometry,—views which have been set forth in their most general form by Plücker in his earlier works.

Those who have penetrated into the spirit of Plücker's works find nothing essentially new in the idea that one may employ as element in the geometry of space any curve involving three parameters. Since no one, as far as I know, has put this suggestion into effect, the reason is probably that no resultant advantages have seemed likely.

I was led to a general study of this theory by discovering that through a very remarkable representation[1] the theory of principal tangent curves can be led back to the theory of curves of curvature.

Following Plücker's plan I shall discuss the system of equations

$$F_1(x, y, z, X, Y, Z) = 0, F_2(x, y, z, X, Y, Z) = 0$$

which, in a sense that will be explained later, defines a general reciprocal relation between two spaces. If, as a special case, the two equations are linear with respect to each system of variables, we obtain a representation in which to the points of one space there correspond in the other the lines of a Plücker complex of lines. The simplest one of the class of transformations derived in this manner is the well-known Ampère transformation, which by this method appears in a new light. I am now making a special study of the method of this representation; for on this I base *a fundamental relation between the Plücker line geometry and a space geometry in which the element is the sphere,*—a very important relation, it seems to me.

While occupied with this paper I have been continually exchanging opinions and views with Plücker's pupil, Dr. Felix Klein. To him I am indebted for many of the ideas here expressed; for some of them I may not even be able to give the reference.

Let me also remark that this paper has several points of contact with my works on the imaginaries of plane geometry. The reason for my not bringing out this dependence in the present discussion is partly that this relation is to some extent fortuitous, and partly

[1] [Lie uses the word "afbildning" which literallym eans picturing or imaging. Since these are uncommon forms in English, we shall use the word "representation" which has been used by later students of the theory. Its graphical connotation is inadequate, however, and we shall use the forms "image" and "imaged" for the words "billede" and "afbildes," consistently used by Lie in his earliest memoirs.

The excessive use of italics is as in the original. It has been thought best to follow Lie's usage as serving to show his points of emphasis.]

that I do not wish to deviate from the customary language of mathematics.[1]

PART I

Concerning a New Space Reciprocity

§1

Reciprocity between Two Planes or between Two Spaces

1. The Poncelet-Gergonne theory of reciprocity can be derived for the field of plane geometry from the equation

$$X(a_1x + b_1y + c_1) + Y(a_2x + b_2y + c_2) + (a_3x + b_3y + c_3) = 0 \tag{1}$$

or from the equivalent equation

$$x(a_1X + a_2Y + a_3) + y(b_1X + b_2Y + b_3) + (c_1X + c_2Y + c_3) = 0,$$

provided that (x, y) and (X, Y) are interpreted as Cartesian point coordinates for two planes.

For if we apply the expression *conjugate* to two points (x, y) and (X, Y) whose coordinate values satisfy equation (1), we may say that the points (X, Y) conjugate to a given point (x, y) form a right line which we may interpret as *corresponding* to the given point.

Since all points of a given right line have a common conjugate point in the other plane, their corresponding right lines pass through this common point.

Thus the two planes are imaged, the one on the other, by equation (1) *in such a way that, mutually, to the points of one plane correspond the right lines of the other To the points of a given right line λ correspond the right lines that pass through λ's image point.*

But this is exactly what constitutes the principle of the Poncelet-Gergonne theory of reciprocity.

Now consider in one plane a multilateral whose vertices are $p_1, p_2, \ldots p_n$, and in the other plane the polygon whose sides $S_1, S_2, \ldots S_n$ correspond to these points. From what has been said it follows also that the vertices $S_1S_2, S_2S_3, \ldots S_{n-1}S_n$, of the latter

---

[1] The theories set forth in this memoir have induced Mr. Klein, in a note just made public (Gesellschaft d. Wissensch. zu Göttingen, March 4, 1870), to carry Plücker's ideas one step forward; for he has demonstrated that the Plücker line geometry (or, in my representation, the corresponding sphere geometry) illustrates the metric geometry of four variables.

multilateral are image points of the sides $p_1p_2, p_2p_3, \ldots p_{n-1}p_n$ of the given polygon, and that the two polygons are thus in a reciprocal relation to one another.

By the process of limits we may pass to considering two curves c and C which correspond in such a way that the tangents of one are imaged as the points of the other. Two such curves are said to be reciprocal to one another in respect to equation (1).

2. Plücker[1] has based a generalization of this theory on the interpretation of the general equation

$$F(x, y, X, Y) = 0. \tag{2}$$

The points (X, Y) [or (x, y)] conjugate to a given point (x, y) [or (X, Y)] now form a curve C [or c] which is represented by equation (2), provided (x, y) [or (X, Y)] be taken as parameters while (X, Y) [or (x, y)] be taken as current coordinates.

Thus, by means of equation (2) the two planes are imaged, the one on the other, in such a way that to the two points in one correspond one-to-one the curves of a certain curve net in the other.

Reasoning as before, we see that to the points of a given curve c [or C] there correspond curves C [or c] which pass through the image point of the given curve.

To a polygon of curves $c(c_1, c_2, \ldots c_n)$ correspond n points $P_1, P_2, \ldots P_n$, which lie in pairs on the curves $C(P_1P_2, P_2P_3, \ldots P_{n-1}P_n)$, whose image points are vertices of the given curvilinear polygon. Here also we come at length to the consideration of curves σ and Σ in the two planes, which are so related that to the points of the one correspond the curves c[or C] that envelope the other. This reciprocal relation, however, is generally not complete, for adjoined forms appear, as a rule.

3. Plücker[2] bases the general reciprocity between two spaces on the interpretation of the general equation

$$F(x, y, z, X, Y, Z) = 0.$$

If F is linear with respect to each system of variables, the Poncelet-Gergonne reciprocity between the two spaces is obtained.

In this memoir, especially in Part One of the same, I aim to make a study of a new space reciprocity to be thought of as coordinate with the Plückerian, and defined by the equations

$$F_1(x, y, z, X, Y, Z) = 0,$$
$$F_2(x, y, z, X, Y, Z) = 0,$$

[1] *Analytisch-geometrische Entwickelungen.* T. I. Zweite Abth.

[2] Though I cannot give any reference, I think I am correct in ascribing this reciprocity to Plücker.

where (x, y, z) *and* (X, Y, Z) *are to be interpreted as point coordinates for the two spaces* r *and* R.

§2

A Space Curve Involving Three Parameters May Be Selected as Element for the Geometry of Space

4. A transformation of geometric propositions which is based on the Poncelet-Gergonne or the Plücker reciprocity may be studied from a higher point of view, as was stressed by Gergonne and Plücker. This view-point will be described here, as it applies also to our new reciprocity.

Cartesian analytic geometry translates any geometric theorem into an algebraic one and effects that the geometry of the plane becomes a physical[1] representation of the algebra of two variables and likewise that the geometry of space becomes an interpretation of the algebra of three variable quantities.

Plücker has called our attention to the fact that the Cartesian analytic geometry is encumbered by a two-fold arbitrariness.

Descartes represents a system of values for the variables x and y by a *point* in the plane; as ordinarily expressed, he has *chosen the point as element for the geometry of the plane,* whereas one could with equal validity employ for this purpose the right line or any curve whatsoever depending on two parameters. In respect to the plane we may therefore look upon the transformation based on the Poncelet-Gergonne reciprocity as consisting of changing from the point to the right line as element, and in the same sense the Plücker reciprocity of the plane consists in introducing a curve involving two parameters as element for the geometry of the plane.

Furthermore, Descartes represents a system of quantities (x, y) by *that* point in the plane whose distances from the given axes are equal to x and y; *from an infinite number of possible coordinate systems he has chosen a particular one.*

The progress made by geometry in the 19th century has been made possible largely because this two-fold arbitrariness in the Cartesian analytic geometry has been clearly recognized as such; the next step should be an effort to utilize these truths still further.

5. The new theories advanced in the following pages are based on the fact that *any space curve involving three parameters may be*

[1] [Lie uses the word "sanselig," affecting the senses, material. We could have translated the word with "visual," but that word often refers to graphical representation in analytic geometry.]

selected as element for the geometry of space. If we recall, for example, that the equations for a right line in space contain four essential constants, we readily see that the right lines satisfying one given condition can be employed as elements for a geometry of space which, like our conventional geometry, gives a physical representation of the algebra of three variables.

This, however, causes a certain system of lines,—*a Plücker complex of lines*—to be singled out, and for this reason it is evident that a particular representation of this kind can have only a limited applicability. However, if it is a question of a *study of space relative to a given complex of lines* it may prove very advantageous to choose the right lines of this complex as space element. In *metrical* geometry the infinitely distant imaginary circle and, hence, the right lines intersecting it are singled out, and *we might therefore have some reason a priori to suppose that in dealing with certain metrical problems, it would be advantageous to introduce these right lines as elements.*

It should be noted that when we as an illustration stated that it is possible to choose the right lines of a line complex as space element, then this is something different—something more particular, if you please—from the ideas that are the basis of Plücker's last work, *Neue Geometrie des Raumes, gegründet auf die Betrachtung der geraden Linie als Raum-Element.* Early in his studies Plücker had observed that it is possible to set up a representation for an algebra which comprises any number of variables by introducing as element a figure depending on the necessary number of parameters. He emphasized[1] particuarly that since the space line has four coordinates, one may, by choosing it as space element, obtain a geometry for which space has *four* dimensions.

§3

The Curve Complex. A New Geometric Interpretation of Partial Differential Equations of the First Order. The Principal Tangent Curves of a Line Complex.

6. Plücker employs the expression *line complex* to designate the assemblage of right lines which satisfy one given condition and which therefore depend on *three* undetermined parameters.

[1] *Geometrie des Raumes.* Art. 258. (1846.)

Analogously I shall define a *curve complex* to mean any system of space curves c, whose equations

$$f_1(x, y, z, a, b, c) = 0, f_2(x, y, z, a, b, c) = 0 \tag{3}$$

contain *three essential constants.*

By differentiating (3) with respect to x, y, z and eliminating a, b, c between the two new and the original equations, we obtain a result in the form

$$f(x, y, z, dx, dy, dz) = 0 \tag{4}$$

If we interpret x, y, z as parameters and dx, dy, dz as direction cosines, then by equation (4) every point in space will be associated with a cone, namely the assemblage of the tangents to the complex curves c which pass through the point in question. These cones I shall call *elementary complex cones.* I shall also use the expression *elementary complex directions* to indicate any line element (dx, dy, dz) belonging to a complex curve c. *The assemblage of the elementary complex directions corresponding to a point form the elementary complex cone associated with the point.*

To a given system (3), or, if we choose, to a given complex of curves there corresponds a definite equation $f = 0$; *but an equation* $f = 0$ *may, on the other hand, be derived from an infinite number of systems* (3).

For, if we choose any relation of the form

$$\psi(x, y, z, dx, dy, dz, \alpha) = 0,$$

where α denotes a constant, and represent by

$$\varphi_1(x, y, z, \alpha, \beta, \gamma) = 0, \varphi_2(X, Y, Z, \alpha, \beta, \gamma) = 0$$

the integral of the simultaneous system

$$f = 0, \psi = 0,$$

then it is clear that if we differentiate $\varphi_1 = 0$, $\varphi_2 = 0$ with respect to x, y, z and eliminate α, β, γ we obtain the result $f = 0$.

Every curve of this new complex

$$\varphi_1 = 0, \varphi_2 = 0$$

is enveloped by the curves c, inasmuch as its elements are severally complex-directions.

7. According to Monge a partial differential equation of the first order in x, y, z is equivalent to the following problem: To find the most general surface which at every one of its points

touches a cone associated with that point, the general equation of the cone in plane coordinates being represented by the given partial differential equation.

Lagrange and Monge have reduced this problem to the determination of a certain complex of curves, the so-called *characteristic curves*, by proving that if we unite into one surface a family of characteristic curves, each of which intersects the curve immediately preceding, an integral surface is always formed.

Note that the equation

$$f(x, y, z, dx, dy, dz) = 0,$$

determined by the characteristic curves as stated above, is of equal value with the partial differential equation itself, for both of these equations are the analytical definition of the same triple infinity of cones.

8. *A more general geometric interpretation of partial differential equations of the first order in x, y, z may be obtained by showing that the problem of finding the most general surface which at every one of its points has a three-point contact with a curve of a given curve complex finds its analytical expression in a partial differential equation of the first order; granted, that the curve in question does not lie wholly on the surface. Furthermore, if $f(x, y, z, dx, dy, dz) = 0$ is the equation determined by the characteristic curves, then will every curve complex whose equations satisfy $f = 0$ stand in the given geometric relation to the given partial differential equation.*

Consider that we have given a complex of curves c which satisfy the equation $f = 0$ and express analytically the requirement that the surface $z = F(x, y)$ have a three-point contact with a curve c at every one of its points, *without excluding the possibility of even a closer contact.* This gives us for the determination of z a partial differential equation of the second order ($\delta_2 = 0$).[1] But every surface generated by an infinity of c's obviously satisfies the equation $\delta_2 = 0$ and therefore its general integral includes two arbitrary functions. By means of analytical considerations that are in essence very simple, though formally somewhat extensive, I wish to prove that the first order differential equation $\delta_1 = 0$, which corresponds to $f = 0$, satisfies $\delta_2 = 0$. Obviously $\delta_1 = 0$ is not, in general, contained in the general integral mentioned; consequently $\delta_1 = 0$ is a *singular* integral of $\delta_2 = 0$.

[1] $\delta_2 = 0$ has the form $A(rt - s^2) + Br + Cs + Dt + E = 0$. Compare a paper by Boole in Crelle's *Journal*. Vol. 61.

The equation $f(x, y, z, dx, dy, dz) = 0$ gives by differentiation,

$$f'_x dx + f'_y dy + f'_z dz + f'_{dx} d^2x + f'_{dy} d^2y + f'_{dz} d^2z = 0, \quad (6)$$

in which $dx, dy, dz, d^2x, d^2y, d^2z$ are considered as belonging to any curve that satisfies $f = 0$. Equation (6) holds, specifically, for the characteristic curves of $\delta_1 = 0$, and if we distinguish these by a subscript, we obtain:

$$f'_{x_1} dx_1 + \ldots f'_{dx_1} d^2x_1 + \ldots = 0$$

Here I shall remark that every curve which touches any of the integral surfaces $U = 0$ of $\delta_1 = 0$, satisfies the equation

$$\frac{dU}{dx}dx + \frac{dU}{dy}dy + \frac{dU}{dz} = 0; \quad (7)$$

and furthermore that every curve which has a three-point contact with $U = 0$ also satisfies the relation:

$$\frac{d^2U}{dx^2}(dx^2) + \ldots \left(\frac{dU}{dx}\right)d^2x \ldots = 0. \quad (8)$$

From this it is seen that every characteristic curve which lies in $U = 0$ satisfies both (7) and (8).

But $U = 0$ touches at every one of its points the associated cone of the system $f = 0$, and therefore the following equations hold:

$$d'_{dx} = \rho\frac{dU}{dx}, f'_{dy} = \rho\frac{dU}{dy}, f'_{dz} = \rho\frac{dU}{dz},$$

where ρ indicates an unknown proportionality factor. Thus the subscripted equation (8) is transformed into the following:

$$\rho\left[\frac{d^2U}{dx_1^2}(dx_1)^2 + \ldots\right] + [f'_{dx_1} d^2x_1 + \ldots] = 0.$$

Now we know that

$$f'_{x_1} dx_1 + \ldots + f'_{dx\cdot} d^2x_1 + \ldots = 0$$

Consequently

$$\rho\left[\frac{d^2U}{dx_1^2} + \ldots\right] = f'_{x_1}[dx_1 + \ldots]$$

or, by omitting the now unnecessary subscripts:

$$\rho\left[\frac{d^2U}{dx^2}dx^2 + \ldots\right] = f'_x dx + \ldots$$

Since

$$\rho\left[\frac{dU}{dx}d^2x + \frac{dU}{dy}d^2y + \frac{dU}{dz}d^2z\right] = [f'_{dx}d^2x + \ldots]$$

the following equation holds:

$$\rho\left[\frac{dU}{dx}d^2x + \frac{dU}{dy}d^2y + \frac{dU}{dz}d^2z + \frac{d^2U}{dx^2}(dx)^2 + \ldots\right]$$
$$= f'_x dx + f'_y dy + f'_z dz + f'_{dx}d^2x + f'_{dy}d^2y + f'_{dz}d^2z,$$

whose left and right-hand members vanish simultaneously.

Our exposition shows that every curve which satisfies $f = 0$ and which touches a characteristic curve lying on $U = 0$ has a three-point contact with this surface; consequently, $\delta_1 = 0$ is a singular integral of $\delta_2 = 0$.

Now we shall prove that $\delta_2 = 0$ has no other singular integral.

For, let every point on an integral surface I of $\delta_2 = 0$ have associated with it a direction, namely, the tangent of the corresponding c of three-point contact. Assuming that I is not generated by a family of c's, there will pass through every point of I two coincident curves c, both tangent to the surface at the point in question. But I is consequently touched at each of its points by the corresponding elementary complex cone; I satisfies the equation $\delta_1 = 0$.

9. *Corollary. The determination of the most general surface which at each of its points has a principal tangent not lying on the surface belonging to a given line complex depends on the solution of a first-order partial differential equation whose characteristic curves are enveloped by the lines of the complex. In this case these curves appear as principal tangent curves on the integral surfaces.*

We shall give an independent geometric proof for this corollary.

The partial differential equation whose characteristic curves are enveloped by the lines of a given line complex is, according to Monge's theory, the analytical expression of the following problem: To find the most general surface which at every one of its points touches the complex cone corresponding to the point. But if the tangents to a curve belong to a line complex, then is the osculation plane of the same the tangent plane of the corresponding complex cone. Thus the osculation planes of our characteristic curves are tangent planes for all integral surfaces that contain these curves. This might require a few additional words of explanation, but it would be largely a repetition of what has been said before.

Accordingly, every complex of lines determines a complex of curves which are enveloped by the lines of the line complex and which possesses this property: they are principal tangent curves on every surface generated by a system of these curves, every curve intersecting the one immediately preceding. *This complex of curves we shall call the principal tangent curves of the line complex.*

I am indebted to Mr. Klein for the statement that the congruence of right lines which Plücker calls the *singular lines* of a line complex belongs to the aforementioned complex of curves. If the given complex is formed by the tangents of a surface [or by the right lines which cut a curve], then are severally the lines of this line complex singular lines and, hence, principal tangent curves.

§4

The Equations $F_1(x, y, z, X, Y, Z) = 0$, $F_2(x, y, z, X, Y, Z) = 0$, Determine a Reciprocity Between Two Spaces.[1]

10. We shall now begin the study of the space reciprocity determined by the equations

$$\begin{aligned} F_1(x, y, z, X, Y, Z) &= 0 \\ F_2(x, y, z, X, Y, Z) &= 0 \end{aligned} \tag{9}$$

where (x, y, z) and (X, Y, Z) are considered point coordinates in two spaces r and R.[2]

If we use the expression *conjugate* about two points the values of whose coordinates (x, y, z) and (X, Y, Z) fulfill the relations (9), we may say that the points (X, Y, Z) conjugate to a given point (x, y, z) form a curve C which is represented by (9), provided x, y, z are interpreted as parameters and X, Y, Z as current coordinates.

To the points of the space r, therefore, correspond one-to-one the curves C of a certain curve complex in R, and there is likewise in r a complex of curves c holding a similar relation to the points of R.

Thus, by equations (9) *the two spaces are imaged, the one in the other, in such a way that to the points in each of the two spaces there correspond one-to-one the curves of a certain complex in the other. As a point describes a complex curve, the complex curve corresponding*

[1] Compare this article with §1.

[2] Whatever pertains to space r we generally designate by small letters, and whatever refers to space R by capital letters.

to the point will turn[1] *about the image point of the described complex curve.*

11. We may now prove that the equations (9) determine a general reciprocity between figures in the two spaces and, specifically, between curves that are enveloped by the complex curves c and C.

When two curves of one complex have a common point (which obviously is not the case in general), their image points lie on a complex curve. Note, specifically, that two infinitely near complex curves which intersect are imaged as two points whose infinitely small connecting line is an elementary complex direction.

Consider in r a curve σ, enveloped by curves c, and all the curves C which correspond to the points of σ. According to our analysis above, two consecutive C's will intersect, and therefore their aggregate will determine an envelope curve Σ.

It is also evident that as a point moves along Σ, the corresponding c will *envelope* a curve σ^* and it can be shown that σ^* is precisely the original given curve σ.

For, consider on the one hand a curvilinear polygon formed by the complex curves $c_1, c_2, c_3, \ldots c_n$, whose vertices are $c_1c_2, c_2c_3, \ldots c_{n-1}c_n$, and on the other hand the image points $P_1, P_2, \ldots P_n$ of the curves c. Manifestly these lie in pairs $P_1P_2, P_2P_3, \ldots P_{n-1}P_n$ on the complex curves C, namely, on those curves which correspond to the vertices of the given polygon. The new polygon in R and the given polygon are therefore in complete reciprocal relation to one another.

By passing to the limit we obtain in the two spaces curves that are enveloped by the complex curves c *and* C, *and which are so reciprocally related that to the points of one correspond the complex curves which envelope the other.*

Therefore a curve enveloped by complex curves is imaged in a two-fold sense as another curve likewise enveloped by complex curves. We say the latter is *reciprocal* to the given curve relative to the system of equations (9). Notice also that the elementary complex directions (dx, dy, dz), (dX, dY, dZ) arrange themselves in pairs as reciprocals, and thus that two curved lines, tangent to one another and enveloped by complex curves, are imaged in the other space as curves bearing the same relation to one another.

[1] The expression "turn" is in-so-far unfortunate as we, of course, mean a turning accompanied with a change of form.

12. There are other space forms between which equations (9) determine a correspondence, which, however, is not generally a complete reciprocity.

Thus, the points of a given surface f are imaged in R as a double infinity of curves C, that is, a congruence of curves whose focal surface[1] *is F. Similarly, there corresponds to the points of F a congruence of curves c, whose focal surface, as we shall see later, contains f as a reducible part.*

The elementary complex cones whose vertices lie in the surface f intersect the corresponding tangent planes of the surface in n right lines (n designating the order of the complex cones) and determine n elementary complex directions at every point of f. The continuous succession of these directions form a family of curves n-ply covering the plane f. The curves are one and all enveloped by complex curves c. *The geometric locus of the reciprocal curves of this family of curves, or, if we choose, the assemblage of the image points of the c's that are tangent to f, form the focal surface F.*

To prove this we recall that two infinitely near and intersecting curves C are imaged as two points whose infinitely small connecting line is an elementary complex direction. From a point p_0 on the surface f proceeds n complex directions. Hence C_0, the image curve of p_0, is intersected in n points by neighboring C's belonging to the curve congruence discussed above. The intersection points are the n points that correspond to the n complex curves c which touch the surface f at the point p_0. Thus the points of F are the image of the c's that touch f.

Since the position of f in space r is general, a curve c which touches f at some one point will in general not have any other points of contact with the surface. But all these c's form a congruence in which every c touches the focal system in N points,—N indicating the order of the elementary complex cones in R. Therefore, as was stated above, the focal system of the congruence is broken up into f and a surface φ, to which every c is tangent in $(N - 1)$ points.

Accordingly, in order that the correspondence between surfaces in r and R determined by equations (9) shall be a complete reciproc-

[1] In analogy with the terminology applied to congruencies of lines I shall take the focal surface of this congruence of curves to mean the geometrical locus of the intersection points of the infinitely near curves C. If we think of a curve congruence as defined by a linear partial differential equation, then its focal surface is what we ordinarily call the singular integral of the differential equation.

ity, it is necessary and sufficient that both n and N be equal to unity. *The reciprocal relation is generally incomplete, inasmuch as analogous operations on the one hand carry f into F, and on the other hand F into the sum of f and φ.*

The above observations are also valid if f, and consequently F, are surface elements; if f is infinitely small in one direction, the same holds for F.

Finally, consider a curve k not enveloped by complex curves c, together with the surface F, formed by all the C's that correspond to the points of k. The points of a C change into the curves c that pass through the image point of C. Hence, to the points of F correspond the assemblage of curves c intersecting k. *Thus there is a two-fold relation of dependence between k and F.*

The equations (9), which picture the two spaces in one another mutually, accordingly carry given space forms into new ones which hold a reciprocal relation to the given forms and therefore serve to transform geometrical theorems and problems. We shall later make important applications of this principle of transformation to a special form of equations (9).

§5

The Transformation of Partial Differential Equations

13. Legendre[1] was the first to give a general method for transforming, in the language of modern geometry, a partial differential equation in point coordinates x, y, z into a differential equation in plane coordinates t, u, v, or (we might also say) in point coordinates t, u, v for a space related reciprocally to the given space.

In a similar manner, if we introduce the curves c as element for the space r it is possible to transform a partial differential equation in x, y, z into a differential equation in the coordinates X, Y, Z of the new space element. In this we may interpret X, Y, Z as point coordinates for the space R,—an interpretation which will be prominent in our presentation.

Let there be given any partial differential equation of the first order in x, y, z, and all the surfaces ψ which represent its so-called "integral complet," bearing in mind that every other integral surface f may be represented as an envelope of a singly infinite set of ψ's.

[1] Compare Plücker, *Geometrie des Raumes.* (1846.) §2.

Consider, in addition, all surfaces Ψ and Φ in space R that correspond to the surfaces ψ and f. We shall prove that every F is the envelope surface of a singly infinite set of Ψ's, that accordingly the surfaces F satisfy a partial differential equation of the first order for which all Ψ's form an "integral complet."

For, if in r there be given two surfaces possessing a common surface element, they will be imaged in R as surfaces that touch one another; and surfaces possessing infinitely many surface elements in common are changed into surfaces that are tangent along a curve in the manner of the given surface.

Assuming this, let us consider an integral surface f_0 and the singly infinite set of Ψ_0's tangent to f_0 along a characteristic curve; and, finally, the corresponding surfaces F_0 and Ψ. It is clear that F_0 has contact with every Ψ along a curve and, hence, that F_0 is the enveloping surface of all the Ψ_0's.

14. Of special interest is the case there the *partial differential equation* that is transformed is precisely the one *determined by the complex curves c* (cf. §3). In this case it may be shown that the corresponding differential equation in X, Y, Z is broken up into two equations, of which one *is precisely the one that corresponds to the complex curves C.*

For, let there be given an integral surface f of the given differential equation in x, y, z, and all the elementary complex cones corresponding to the points of f. By §4, these cones determine, at every point of f, n complex directions, of which in this case two are coincident; hence the family of curves that are enveloped by the complex curves c and lie on the surface f, which we discussed in §4, is broken up into the characteristic curves of f and a curve system that covers f $(n - 2)$-fold.

Thus the curve congruence in R corresponding to the points of f has a focal system which is separated into two surfaces, of which one, which we shall call Φ, is tangent to every c at two coincident points, while to the other there are $(n - 2)$ points of contact. *Thus the surfaces Φ satisfy the partial differential equation which, according to the theorem in §3, is determined by the complex curves C.*

Noting that Φ is the geometric locus of the reciprocal curves of the characteristic curves of f, we see that two integral surfaces f_1 and f_2, tangent to one another along a characteristic curve k, are transformed into two surfaces Φ_1 and Φ_2 that are tangent to one another along the *reciprocal* curve of k; for k is enveloped by complex curves c.

The characteristic curves of the two partial differential equations which, according to §3, are determined by the curve complexes c and C, are reciprocal curves relative to the system of equations (9).

15. The proposition just stated gives the following general method for transforming partial differential equations of the first order.

Determine by the usual methods the equation

$$f(x, y, z, dx, dy, dz) = 0$$

which the characteristic curves of the given partial differential equation satisfy. Then select a relation of the form

$$\Psi(x, y, z, dx, dy, dz, X) = 0,$$

where X denotes a constant. Let the simultaneous system

$$f = 0, \Psi = 0,$$

be integrated in the form

$$F_1(x, y, z, X, Y, Z) = 0, F_2(x, y, z, X, Y, Z) = 0,$$

where Y and Z are the constants of integration. By differentiation and elimination we obtain a relation of the form

$$F_3(X, Y, Z, dZ, dY, dZ) = 0,$$

which we interpret to be the equation of the characteristic curves of a partial differential equation

$$F_4\left(X, Y, Z, \frac{dZ}{dX}, \frac{dZ}{dY}\right) = 0.$$

Our former discussions show that $F_4 = 0$, derived from $F_3 = 0$ by the usual processes, and the given partial differential equation are mutually dependent in such a manner that if one is integrable, so is the other.

From this we may draw general conclusions concerning the reducing to lower degree of first order partial differential equations defined by a complex of curves of a given order. Every first-order partial differential equation defined by a line complex (§3), for example, may be transformed into a partial differential equation of the second degree.[1]

We may likewise transform every partial differential equation defined by a complex of conics into a differential equation of degree 30.[2]

---

[1] This reduction depends on the fact that every line of a line congruence touches the focal system in two points. (§4, 12.)

[2] The number 30 results from the product of 6 by $(6 - 1)$; 6 is the number of points in which the focal system of a congruence of conics has contact with each conic.

§6

Concerning the Most General Transformation Which Change Surfaces Mutually Tangent into Similarly Situated Surfaces

16. In the study of partial differential equations an important role is played by transformations expressible in the form

$$X = F_1(x, y, z, p, q),\ Y = F_2(x, y, z, p, q),$$
$$Z = F_3(x, y, z, p, q),$$

As usual, p and q indicate the partial derivatives $\frac{dz}{dx}, \frac{dz}{dy}$; P and Q likewise stand for $\frac{dZ}{dX}$ and $\frac{dZ}{dY}$.

In the following we shall consider the case[1] where the functions F_1, F_2, and F_3 are chosen such that P and Q also depend only on x, y, z, p, q. Thus:

$$P = F_4(x, y, z, p, q);\ Q = F_5(x, y, z, p, q).$$

Assuming that no relation between X, Y, Z, P, Q can be derived from the above five equations, we shall show that each of the quantities x, y, z, p, q are also expressible as functions of X, Y, Z, P, Q.

If we think of x, y, z and X, Y, Z as point coordinates for r and R, we may say that by a transformation of this kind there is defined a *correspondence between the surface elements of the two spaces,—in fact, the most general correspondence.* We shall show that *these transformations divide into two distinct, coordinate classes, of which one corresponds to the Plücker reciprocity, while the other corresponds to the reciprocity which I have set up in this memoir.*

Eliminating p, q, P, and Q in the five equations

$$X = F_1, \quad Y = F_2, \quad Z = F_3, \quad P = F_4, \quad Q = F_5$$

two essentially different results may come about. We shall either obtain only one equation in x, y, z, X, Y, Z, or there will be two relations obtaining among the quantities. (The existence of *three* mutually independent equations involving the point coordinates of the two spaces assumes that the transformation in question is a *point* transformation.)

But we know that the equation $F(x, y, z, X, Y, Z) = 0$ *always* defines a reciprocal correspondence between the surface elements

---

[1] Cf. Du Bois-Reymond, *Partielle Differential Gleichungen.* §§75–81.

of the two spaces. I have likewise shown in the preceding pages that the system

$$F_1(x, y, z, X, Y, Z) = 0, \quad F_2(x, y, z, X, Y, Z) = 0$$

always determines a transformation which changes mutually tangent surfaces into similarly situated surfaces.

My statement is therefore proved.

Let me at this time call attention to a remarkable property of these transformations: they change any differential equation of the form

$$A(rt - s^2) + Br + Cs + Dt + E = 0,$$

in which A, B, C, D are dependent only on x, y, z, p, q, into an equation of the same form. Consequently, if the given equation has a general first integral, so does the resulting equation (Cf. Boole's paper in Crelle's *Journal*, Vol. 61).

PART II

The Plücker Line Geometry May Be Transformed into A Sphere Geometry

§7

The Two Curve Complexes are Line Complexes

17. Let us assume that these equations, which image the two spaces in one another, are linear in each system of variables:

$$(10)\quad \begin{cases} 0 = X(a_1x + b_1y + c_1z + d_1) + Y(a_2x + b_2y + c_2z + d_2) \\ \qquad + Z(a_3x + b_3y + c_3z + d_3) + (a_4 + \ldots) \\ 0 = X(\alpha_1x + \beta_1y + \gamma_1z + \delta_1) + Y(\alpha_2x + \beta_2y + \gamma_2z + \delta_2) + \\ \qquad Z(\alpha_3x + \beta_3y + y_3z + \delta_3) + (\alpha_4x + \beta_4y + \gamma_4z + \delta_4). \end{cases}$$

Then clearly the points of the other space which are conjugate to a given point will form a right line. The two curve complexes are Plücker line complexes.[1] It follows that the equations (10) define a correspondence between r and R which possesses the following characteristic properties:

(*a*) *To the points in each space correspond one-to-one the lines of a line complex in the other.*

[1] Regarding the theory of line complexes I assume the reader's acquaintance with these two works: Plücker, *Neue Geometrie des Raumes, gegründet auf*, etc. . . . (1868–69); Klein, "Zur Theorie der Complexe," *Math. Annalen*, Vol. II.

(b) As a point describes a complex line, the corresponding line in the other space turns about the image point of the described line.

(c) Curves enveloped by the lines of the two complexes arrange themselves in pairs, as reciprocals, in such a way that the tangents of each one correspond to the points of the other.

(d) With a surface f in space r there is associated in a two-fold sense a surface F in R. On the one hand F is the focal surface of the line congruence of which f is the image; on the other, the points of F correspond to those tangents of f which belong to the line complex in r.

(e) On f and F all curves arrange themselves as pairs of conjugates in such a way that to the points of a curve on f [or F] corresponds in the other space a line surface which contains the conjugate curve and is tangent to F or f along the curve.

(f) To a curve on f enveloped by the lines of the line complex there corresponds conjugately a curve on F also enveloped by complex lines, and these curves are reciprocal in the sense defined in (c).

Each of the equations (10) determine an anharmonic correspondence between points and planes in the two spaces. Consequently each of our line complexes may be defined as the aggregate of the lines of intersection of planes in anharmonic relation, or as the connecting lines of points in anharmonic relation. But according to Reye the second-degree complex thus defined is identical with a certain line system discussed by Binet. Binet was the first to look upon this system as the aggregate of the stationary axes of revolution of a material body. It has since been studied by several mathematicians, notably Chasles and Reye.

If we particularize the constants in equations (10), we either give the two complexes a special position or we particularize the complexes themselves. As to the special positions complexes assume, they may, for example, coincide; and Mr. Reye has discussed this case in his *Geometrie der Lage* (1868), Part Second, where he also gives the propositions stated in (*a*) and (*b*). As regards the particularized complexes, I shall not enter into a discussion of all the possible special cases, but will emphasize two of the most important degenerations:[1]

[1] Lie, "Repräsentation der Imaginaeren," in the proceedings of the Christiania Academy of Sciences (Christiania Videnskabs-Selskab) for February and August, 1869. The space representation there discussed in §§17 and 27–29 is identical with the one discussed here. In §25 I emphasize expressly the first of the two degenerations discussed here.

(1) Both complexes may be *special* and *linear.* This case gives us the well-known transformation of Ampère. We may therefore consider this transformation as based on our introducing as space element the assemblage of right lines intersecting a given line, instead of the point.

(2) One complex may degenerate into the assemblage of right lines that intersect a given conic. In that case the other complex will be a general linear complex. I may mention here that Mr. Noether (*Götting. Nachr.*, 1869) has, on occasions, given a representation of the linear complex in a point space which is identical with the one under discussion. But the conception that *every* space contains a complex whose lines are imaged as the points of the other space, which is fundamental for our purpose, is not touched upon in Mr. Noether's brief presentation.—This is the degeneration of which we shall make a study in the following article. We assume that the fundamental conic is the infinitely distant imaginary circle.

18. We have seen that the two curve complexes are line complexes if the equations of representation are linear in each system of variables. This leads us to investigate whether this sufficient condition is necessary.

If one complex is a general line complex, the elementary complex cones of the corresponding curve complex must be resolved into cones of the second degree. The proof (cf. §4, 12) of this lies in the fact that the lines of a *line* congruence touch the focal surface in *two* points. If one complex is a special line complex, then the elementary complex cones of the corresponding curve complex in the other space will resolve into plane sheaves.

Thus, if both complexes are to be line complexes the elementary complex cones of both spaces must be resolved into second and first degree cones. But if the cones of a line complex may be continually resolved, the complex is itself reducible.[1] We have therefore proved that if two line complexes are imaged in one another as described in the previous article, it follows that either both are of the second degree, or one is a special complex of the second degree and the other linear, or they are both special linear complexes. All three cases are represented by equations (10), and we shall indicate how one may know *that equations* (10) *define the most general representation of two line complexes upon one another.*

---

[1] I know of no proof for this assertion, but I have been told that it is reliable. However, the conclusions based on it are not essential for what follows.

If both complexes are of the second degree it can be shown that the surface of singularity can not be a *curved* surface.

For through each point of this surface there pass two plane sheaves whose lines are imaged in the other space as the points of one right line. From this follows that all of the lines of one sheaf correspond to one and the same point in the other space.

But the assemblage of lines which have not independent images cannot form a complex; they can only, at best, form a congruence or a number of congruences. Since, however, the assemblage of plane sheaves of rays which proceed from each of all the points of a *curved* surface of necessity forms a complex, our assertion that the surface of singularity cannot be a *curved* surface is proved.

If two complexes of the second degree are imaged upon one another—in which case none of them may be a special complex—, the surface of singularity for each will consist of planes, and consequently both line systems are of the kind first studied by Binet.

If a second-degree complex and a linear complex are imaged upon one another, one might in advance conceive of two possible cases: (1) the second degree complex might be formed by all the lines intersecting a conic,—and such a case does exist, according to the above discussion; (2) the second degree complex might consist of all the tangents to a second degree surface. Through considerations having something in common with those I shall use in §12 I have shown that this case does not exist. For if it did, I might deduce, from the fact that a linear complex can be changed into itself by a triple infinity of linear, inter-permutable transformations, that the same would hold for the second degree surface. Which, however, is not so.

§8

Reciprocity between a Linear Complex and the Assemblage of Right Lines Which Intersect the Infinitely Distant Imaginary Circle

19. In what follows we shall make a closer study of the system of equations:

$$\left.\begin{aligned} -\frac{\lambda}{2B}Zz &= x - \frac{1}{2A}(X + iY) \\ \frac{1}{2B}(X - iY)z &= y - \frac{1}{2\lambda A}Z, \end{aligned}\right| \quad i = \sqrt{-1} \tag{11}$$

This is linear in respect to both systems of variables and therefore, according to §7, it determines a correspondence between two line complexes. We shall first derive the equations of these complexes in Plücker line coordinates.

Plücker gives the equations of the right line in the form

$$rz = x - \rho, \qquad sz = y - \sigma,$$

where the five quantities r, ρ, s, σ, $(r\sigma - s\rho)$ are considered line coordinates. Therefore, if we regard X, Y, Z as parameters, equations (11) represent the system of right lines whose coordinates satisfy these relations:

$$r = -\frac{\lambda}{2B}Z, \qquad \rho = \frac{1}{2A}(X + iY),$$

$$s = \frac{1}{2B}(X - iY), \qquad \sigma = \frac{1}{2\lambda A}Z.$$

These by the elimination of X, Y, Z, give as the equation of our complex

$$\lambda^2 A\sigma + Br = 0. \tag{12}$$

Thus the line complex in the space r is a linear complex. It is, furthermore, a general linear complex and contains, as we notice, the infinitely distant right line of the xy-plane.

To determine the line complex in R we replace the system (11) by the equivalent equations

$$\left(\frac{\lambda A}{2B}z - \frac{B}{2\lambda Az}\right)Z = X - \left(Ax + B\frac{y}{z}\right),$$

$$\frac{1}{i}\left(\frac{\lambda A}{2B}Z + \frac{B}{2\lambda Az}\right)Z = Y - \frac{1}{i}\left(Ax - B\frac{y}{z}\right).$$

Comparing these with the equations of the right line in R,

$$RZ = X - P,\ SZ = Y - \Sigma, \tag{13}$$

we have

$$R = \frac{\lambda A}{2B}z - \frac{B}{2\lambda Az}, \qquad P = Ax + B\frac{y}{z},$$

$$S = \frac{1}{i}\left(\frac{\lambda A}{2B}z + \frac{B}{2\lambda Az}\right), \qquad \Sigma = \frac{1}{i}\left(Ax - B\frac{y}{z}\right).$$

The equation of the line complex in R is then found to be

$$R^2 + S^2 + 1 = 0. \tag{14}$$

But by (13),

$$R = \frac{dX}{dZ}, \qquad S = \frac{dY}{dZ},$$

and consequently we may write (14) in the form

$$dX^2 + dY^2 + dZ^2 = 0. \tag{15}$$

From which we see that the line complex in R is formed by the imaginary right lines whose length equals zero, or, if we choose, by the lines which intersect the infinitely distant imaginary circle.

By equations (11) *the two spaces are imaged, the one in the other, in such a way that to the points of r there correspond in R the imaginary right lines whose length is zero, while the points of R are imaged as the lines of the linear complex* (12).

It should be noted that as a point moves along a line of this linear complex, the corresponding right line in R describes an infinitesimal sphere,—a point sphere.

20. According to the general theory of reciprocal curves developed in §4, if we know a curve whose tangents belong to one of our line complexes, it is possible to find by simple operations the image curve that is enveloped by the lines of the other complex. Lagrange made a study of the general determination of space curves whose length is equal to zero, whose tangents therefore possess the same property. He found the general equation of these curves. Therefore, according to the analysis above, *it is also possible to set up general formulas for the curves whose tangents belong to a linear complex.*

So as not to digress from our aim we shall refrain from taking up in detail the simple geometric relations that exist between the reciprocal curves of the two spaces.[1]

We must now somewhat modify our previous observations concerning the correspondence between surfaces in the two spaces, inasmuch as all the congruencies of right lines which intersect the infinitely distant circle possess a common focal curve—namely, the circle itself—, and inasmuch as the right lines of a line congruence touch the focal surface at only two points.

For let there be given a surface F in R and let f be the geometric locus of the points in r that correspond to the tangents to F of

[1] If the given curve of length zero has a cusp, the corresponding curve in the linear complex has a stationary tangent. In general *stationary tangents* appear as *ordinary singularities*, if curves are regarded as formed by lines, that is, as enveloped by lines of a *given* line complex.

length zero. Then, conversely, F is also the *complete* geometric locus of the image points of the right lines in the linear complex (12) which are tangent to f.

On the other hand, if we have given a surface φ in a general position in r, the instance is like the ordinary case; for then the right lines of the linear complex (12) which touch φ also envelope another surface ψ, the so-called reciprocal polar of φ relative to (12).

This system of lines is imaged in R as a surface Φ, which clearly is the focal surface for two congruences,—one being the assemblage of right lines of length zero which correspond to the points of φ, and the other, the assemblage of the lines having the same relation to the points of ψ.

The tangents of length zero of Φ consequently resolve into two systems; or, we may say, the geodetic curves of length zero of Φ form two distinct families.

In passing I wish to remark that the determination of the *curves which are enveloped by the right lines of a congruence belonging to a linear complex may be reduced, according to our general theory, to finding on the image surface F the geodetic curves whose length is zero.* For these curves are mutually reciprocal (17, f) relative to the system (11).

21. Later we shall find use one or two times for the following two propositions:

a. A surface F of the nth order, which includes the infinitely distant imaginary circle as a p-fold line, is the image of a congruence whose order and, consequently, whose class is $(n - p)$.[1]

For, an imaginary line of zero length intersects F in $(n - p)$ points of the finite space; hence there are always $(n - p)$ lines in the image congruence which pass through a given point, or which lie in a given plane in space r.

b. A curve C of the nth order which intersects the infinitely distant circle in p points is imaged in r as a line surface of order $(2n - p)$.

For, a right line of the linear complex (12) intersects this line surface in as many points, numerically, as there are common points (not infinitely distant) between the curve C and an infinitesimal sphere.

---

[1] Let me here state a proposition which is well-known to every mathematician who works with line geometry, but which is not stated explicitly anywhere, as far as I know: *For a congruence belonging to a linear complex, the order is always numerically equal to the class.*

§9

The Plücker Line Geometry May Be Transformed into a Sphere Geometry

22. In this section we shall give the basis for a *fundamental relation that exists between the Plücker line geometry and a geometry whose element is the sphere.*

For equations (11) transform the right lines of space r into the spheres of space R, and in a two-fold sense (12).

On the one hand the right lines of the complex (12) which intersect a given line l_1, and hence also its reciprocal polar l_2 relative to (12), are imaged as the points of a sphere, according to the proposition in (21, b); on the other hand the points of l_1 and l_2 are changed into the right line generatrices of this sphere.

We arrive at the relations that obtain between the line coordinates of l_1 and l_2 and the coordinate of the center X', Y', Z', and the radius H' of the image sphere, by the following analytic observations:

Let the equations of the line l_1 [or l_2] be

$$rz = x - \rho, \qquad sz = y - \sigma.$$

Also recall that the right lines of the linear complex (12) may be represented by the equations

$$-\frac{\lambda}{2B}Zz = x - \frac{1}{2A}(X + iY)$$
$$\frac{1}{2B}(X - iY)z = y - \frac{1}{2\lambda A}Z.$$

It is clear that if we eliminate x, y, z between these four equations we have the relation which expresses the condition that the right lines intersect l_1. By so doing we arrive at the following relation between the parameters X, Y, Z of these lines, or, if we choose, between the coordinates of the image points:

$$\left[Z - \left(A\sigma\lambda - \frac{Br}{\lambda}\right)\right]^2 + [X - (A\rho + Bs)]^2 + [Y - i(Bs - A\rho)]^2$$
$$= \left[A\lambda\sigma + \frac{B}{\lambda}r\right]^2. \qquad (16)$$

The immediate interpretation of this equation confirms the above statements and gives, in addition, the following formulas:

$$X' = A\rho + Bs, \qquad iY' = A\rho - Bs,$$
$$Z' = \lambda A\sigma - \frac{B}{\lambda}r, \quad \pm H' = \lambda A\sigma + \frac{B}{\lambda}r, \qquad (17)$$

or the equivalent formulas:

$$\begin{aligned} \rho &= \frac{1}{2A}(X' + iY'), & s &= \frac{1}{2B}(X' - iY'), \\ \sigma &= \frac{1}{2\lambda A}(Z' \pm H'), & r &= -\frac{\lambda}{2B}(Z' \mp H'). \end{aligned} \tag{18}$$

(We may without loss omit the primes on the *sphere coordinates* X', Y', Z', H', since, in our conception, the points of space R are spheres of radius zero.)

Formulas (17) *and* (18) *show that a right line in r is imaged as a uniquely defined sphere in R, while to a given sphere there correspond in r two lines*

$$(X, Y, Z, + H) \text{ and } (X, Y, Z, - H),$$

which are reciprocal polars relative to the linear complex

$$H = 0 = \lambda A\sigma + \frac{B}{\lambda}r. \tag{12}$$

If H is set equal to zero, formulas (17) and (18) express clearly that the right lines of the complex (12) and the point spheres of space R are of one set in a one-to-one relation.

A plane—that is, a sphere with infinitely large radius—is imaged as two right lines (l_1 and l_2) which intersect the infinitely distant right line of the xy-plane. It follows that the points of l_1 and l_2 are the images of the imaginary lines in the given plane which pass through its infinitely distant circle points.

As a particular case we note that to a plane which touches the infinitely distant imaginary circle there corresponds a line of the complex $H = 0$ parallel to the xy-plane.

23. *Two intersecting right lines l_1 and λ_1, are imaged as spheres in a position of tangency.*

For the polars of l_1 and λ_1 relative to $H = 0$ also intersect one another and consequently the spheres have two common generatrices. But second-degree surfaces whose curves of intersection consist of a conic and two right lines touch one another in three points, the double points of the curve of the section. The image spheres of l_1 and λ_1, therefore, have three points of contact of which two, imaginary and infinitely distant, in common parlance, do not enter into our discussion.

The analytic proof of our theorem follows.

The condition that the two right lines

$$\begin{aligned} r_1 z &= x - \rho_1, & r_2 z &= x - \rho_2 \\ s_1 z &= y - \sigma_1, & s_2 z &= y - \sigma_2 \end{aligned}$$

intersect is expressed by the equation

$$(r_1 - r_2)(\sigma_1 - \sigma_2) - (\rho_1 - \rho_2)(s_1 - s_2) = 0.$$

This, by aid of (18), gives

$$(X_1 - X_2)^2 + (Y_1 - Y_2)^2 + (Z_1 - Z_2)^2 + (iH_1 - iH_2)^2 = 0,$$

which proves our proposition.

Our theorem shows that the assemblage of right lines which intersect a given line is imaged as the totality of all the spheres which touch a given sphere. Consequently *the image of the special linear complex is known.*

Conversely, to two spheres that are tangent to one another there correspond two pairs of lines so situated that every line in one pair intersects a line in the other.

24. *The representation*[1] *of the general linear complex.* The general linear complex is represented by the equation

$$(r\sigma - \rho s) + mr + n\sigma + p\rho + qs + t = 0. \tag{19}$$

Equations (18) and (19) give us, as the equation of the corresponding "linear complex of spheres"

$$(X^2 + Y^2 + Z^2 - H^2) + MX + NY + PZ + QH + T = 0.^2$$

In this equation M, N, P, Q, T denote constants that depend upon m, n, p, q, t, *while* X, Y, Z, H *are understood to be (non-homogeneous) sphere coordinates.*

It is easy to see that the last equation determines all the spheres that intersect at a constant the image sphere of the linear congruence common to the complexes (19) and $H = 0$.

If the simultaneous invariant of these complexes is equal to zero, or if, to use Klein's expression, the two complexes are in involution, then the constant angle is a right angle.

To spheres that intersect a given sphere at a constant angle there correspond in r the right lines of two linear complexes which are reciprocal polars relative to $H = 0$.

We note particularly that the spheres which intersect a given sphere orthogonally are imaged as the right lines of a linear complex in involution with $H = 0$.

---

[1] [Lie uses the word "afbildning," meaning, literally, picture or image.]

[2] This equation may be put in the form

$$(X - X_0)^2 + (Y - Y_0)^2 + (Z - Z_0)^2 + (iH - iH_0)^2 = C^2,$$

where X_0, Y_0, Z_0, H_0, C_0 are understood to be non-homogeneous coordinates of the linear complex. Mr. Klein has called to my attention the fact that the sphere (X_0, Y_0, Z_0, H_0) is the image of the axis of this linear complex.

Let there be given a linear complex whose equation is of the form

$$ar + bs + c\rho + d\sigma + e = 0. \tag{20}$$

The corresponding relation between X, Y, Z, H is also linear, and hence the linear sphere complex is formed by all the *spheres which intersect a given plane at a given constant angle.*

This might also have been deduced from the fact that the complex (20) contains the infinitely distant right line of the xy-plane, and that therefore the congruence common to it and $H = 0$ possesses directrices that intersect this line.

If the complexes (20) and $H = 0$ are in involution, then the lines of (20) are imaged as the totality of spheres that intersect a given plane orthogonally, or, what amounts to the same thing, as the spheres whose centers lie in a given plane.

The four complexes

$$X = 0 = A\rho + Bs, \qquad Z = 0 = \lambda A\sigma - \frac{B}{\lambda}r,$$
$$iY = 0 = A\rho - Bs, \qquad H = 0 = \lambda A\sigma + \frac{B}{\lambda}r,$$

are obviously in involution by pairs. They also contain as a common line the infinitely distant line of the xy-plane.

Thus, the special linear complex (*Constant* = 0), *formed by all the lines parallel to the xy-plane, in conjunction with the four general linear complexes* $X = 0$, $Y = 0$, $Z = 0$, $H = 0$, *forms a system which we may regard as a degeneration of Mr. Klein's six fundamental complexes. In analogy with our use of* X, Y, Z, H *as non-homogeneous coordinates for a geometry of four dimensions, with the sphere as element introduced above, we may also use these quantities as non-homogeneous line coordinates.*

It is interesting to notice that the linear complexes whose equation is

$$H = \lambda A\sigma + \frac{B}{\lambda}r = \text{constant},$$

and which, according to the form of the equation, are tangent to one another in a special linear congruence, the directrices of which unite in the infinitely distant line of the xy-plane, are imaged as a family of sphere complexes characterized by the property that all their spheres have equal radii.

25. *Various Representations.* A surface f and all its tangent lines at a given point are imaged as a surface F and all the spheres that are tangent to it at a given point.

A line on f is imaged as a sphere which is tangent to F along a curve

If f is a line surface, then F is a sphere envelope,—a tubular surface.

If, particularly, f is a second degree surface and, hence, contains two systems of right line generatrices, then we may interpret F as a sphere envelope in two ways. It is clear that in this manner we *obtain the most general surface possessing this property* (*the cyclide*).

A developable surface changes into the envelope surface of a family of spheres in which two consecutive spheres are tangent to one another throughout,—that is, into an imaginary line surface whose generatrices intersect the infinitely distant imaginary circle. These line surfaces, we know, are precisely the ones characterized by Monge as possessing only one system of curves of curvature.

26. An immediate consequence of Plücker's conception is that if $l_1 = 0$ and $l_2 = 0$ are the equations of two linear complexes, then the equation $l_1 + \mu l_2 = 0$, where μ is a parameter, represents a family of linear complexes that include a common linear congruence. The principle of representation which we employ transforms this theorem into the following:

The spheres K which intersect two given spheres S_1 and S_2 at given angles V_1 and V_2 hold the same relation to infinitely many spheres S. Corresponding to the two directrices of the line congruence are two spheres S, to which all the spheres K are tangent.

The variable line complex $l_1 + \mu l_2 = 0$ intersects the complex $H = 0$ in a linear congruence whose directrices describe a second-degree surface, namely, the section of the three complexes $l_1 = 0$, $l_2 = 0$, $H = 0$. Consequently the spheres S envelope a cyclide. In this instance the cyclide degenerates into a circle along which the different spheres S intersect.

We wish to call attention to the fact that our sphere representation enables us to derive from intersecting discontinuous groups of lines corresponding groups of spheres, and conversely. As an instance, we may apply the well-known theory concerning the twenty-seven right lines of a third-degree surface to prove the existence of groups of twenty-seven spheres, of which each one is tangent to ten of the others.

Conversely, piles of spheres present lines of a linear complex arranged in peculiar, discontinuous arrays.

§10

Transforming Problems Concerning Spheres into Problems of Lines

27. In this section we shall solve a few simple and familiar problems concerning spheres by considering the corresponding line problems that result from our principle of transformation.

Problem I.—How many spheres are tangent to four given spheres?

The four spheres are transformed into four pairs of lines (l_1, λ_1), (l_2, λ_2), (l_3, λ_3), (l_4, λ_4). The corresponding problem of lines is, therefore, to find the lines that intersect four lines selected from the eight in such a way that each pair furnishes one line.

Lines l and λ may be arranged in sixteen different groups of four, in such a way that each group contains only one line from each pair; thus:

$$l_1l_2l_3l_4,\ \lambda_1\lambda_2\lambda_3\lambda_4$$
$$l_1l_2l_3\lambda_4,\ \lambda_1\lambda_2\lambda_3l_4$$
$$\cdots\cdots\cdots\cdots$$
$$\cdots\cdots\cdots\cdots$$

But these sixteen groups are also formed in pairs by lines that are reciprocal polars in respect to $H = 0$. Consequently the pairs of transversals (t_1, t_2) (τ_1, τ_2) of two related groups are also one another's polars in respect to $H = 0$. The last-mentioned four lines are therefore imaged as *two* spheres, and consequently there exist sixteen spheres arranged in eight pairs, which are tangent to four given spheres.

Problem II.—How many spheres intersect four given spheres at four given angles?

The spheres which intersect a given sphere at a fixed angle are imaged as those right lines of two linear complexes which are mutually reciprocal polars in respect to $H = 0$. We must therefore observe four pairs of complexes, (l_1, λ_1), (l_2, λ_2), (l_3, λ_3), (l_4, λ_4), and the problem now is, to find those lines which belong to four of these complexes and are selected in such a way that one is taken from each pair.

Four linear complexes have two common lines. Therefore, if we follow the same procedure as was used in the preceding problem, we shall obtain as the solution sixteen spheres arranged in eight pairs.

Our problem is simplified if one or more of the given angles are right angles; for then the spheres orthogonal to a given sphere are imaged as the lines of *one* complex, which is in involution with $H = 0$ (Article 24). If all angles are right angles, the question is, how many lines are common to four linear complexes in involu-

tion with $H = 0$. Two such lines are mutually reciprocal polars in respect to $H = 0$, and *consequently there is only one sphere which intersects four given spheres orthogonally.*

Problem III.—To construct the spheres which intersect five given spheres at a fixed angle.

Our principle of transformation changes this problem into the following: To find the linear complexes which contain one line from each of five given pairs $(l_1, \lambda_1) \ldots (l_5, \lambda_5)$.

These ten lines may be arranged in thirty-two different groups of five in a way such that every group contains *one* line of each, thus:

$$(l_1 l_2 l_3 l_4 l_5), \ (\lambda_1 \lambda_2 \lambda_3 \lambda_4 \lambda_5)$$

. . . .

. . . .

Note that these groups are mutually reciprocal polars by pairs in respect to $H = 0$. Every group gives a line complex and in all we obtain thirty-two linear complexes conjugate in pairs. These are imaged as sixteen linear sphere complexes. The sixteen spheres which are severally intersected at a constant angle by the spheres of these systems are the solutions of our problem.

Two groups of lines, as $(l_1, l_2, \lambda_3, \lambda_4, l_5)$ and $(\lambda_1, l_2, \lambda_3, \lambda_4, l_5)$ contain four common lines. It follows that the two corresponding linear complexes intersect in a linear congruence whose directrices d_1 and d_2 are the transversals of these four lines.

But the complex $H = 0$ intersects this congruence along a second degree surface which is the image of a circle, namely, the section circle of two of the spheres wanted, as also of the image spheres of d_1 and d_2. The latter spheres may be defined by saying they are tangent to four out of five given spheres; hence, by the aid of the construction just described, we may determine a number of circles on any of the spheres wanted.

On each of the sixteen spheres which intersect five given spheres at a constant angle we may construct five circles, provided we can construct the spheres that are tangent to four given spheres.

§11

The Relation between the Theory of Curves of Curvature and the Theory of Principal Tangent Curves

28. The transformation discussed in the previous sections acquires a peculiar interest on account of the following, in my opinion, very important theorem:

To the curves of curvature of a given surface F in space R there correspond in space r line surfaces which touch the imaged surface f along principal tangent curves.

The tangents of the surface f are transformed into spheres that touch F, and the thought lies near that *to the principal tangents of f there correspond the principal spheres of F.* This also proves to be the case.

For f is cut by a principal tangent in three coincident points, and this shows that three consecutive generatrices of the image sphere of the principal tangent touch F. But such a sphere cuts F along a curve which has a cusp in the contact point of the two, and this is precisely a characteristic of principal spheres.

Note, furthermore, that the direction of this cusp is tangent to a curve of curvature. It is then seen that two consecutive points of a principal tangent curve on f are imaged as two lines which touch F at consecutive points of the same curve of curvature. Therefore, *to the principal tangent curves of f, considered as formed by points, there correspond imaginary line surfaces that touch F along curves of curvature.*

But curves on f and F arrange themselves in pairs of conjugate curves in such a way (Article 17, e) that the points of one form the image of lines that touch the other surface at points of the conjugate curve. *This proves our theorem.*

The following two illustrations may be regarded as verifications of this proposition.

A sphere in R is the image of a linear congruence, of which the two directrices are to be considered the focal surface. We know that every curve on a sphere is a curve of curvature. Moreover, the directrices appear as principal tangent curves on every line surface belonging to a linear congruence.—An hyperboloid f in space r presents in R a surface which in two ways may be regarded as a sphere envelope. But the line surfaces in the complex $H = 0$ which touch f along its principal tangent curves, that is, along its right line generatrices, are themselves surfaces of the second degree. Consequently *the curves of curvature of the cyclide F are circles.*

An interesting corollary resulting from our theorem is the following:

Kummer's surface of order and class four has algebraic principal tangent curves of order sixteen, and these form the complete contact section between this surface and line surfaces of order eight.

For, Kummer's surface is the focal surface of the general line congruence of order and class two which is imaged (provided it belongs to $H = 0$) as a fourth degree surface containing the infinitely distant circle twice (Article 21, *b*).

Now, Darboux and Moutard[1] have demonstrated that the lines of curvature of the last-mentioned surface are curves of order eight, cutting the infinitely distant imaginary circle in eight points. Hence, these lines are imaged as line surfaces of order eight (Article 21, *b*).

If we recall that the generatrices of these line surfaces are double tangents to the Kummer surface, we shall perceive the correctness of the proposition.[2]

It is clear that the degenerations of the Kummer surface, as, for example, *the wave surface, the Plücker complex plane, the Steiner surface of order four and class three,*[3] *a line surface of the fourth degree, the line surface of the third degree,* also have algebraic principal tangent curves.

29. Mr. Darboux has proved that we can generally determine a line of curvature in finite space on any surface, the curve of contact with the imaginary developable, which simultaneously is circumscribed about the given surface and the infinitely distant imaginary circle.

In consequence of which we can generally point out one principal tangent curve on the focal plane of a congruence of a linear complex, this curve being the geometric locus of points for which the tangent plane is also the plane associated with the linear complex.

For the infinitely small spheres which are tangent to F consist of the points of F and of the above-described imaginary developable. Consequently the right lines of the complex $H = 0$ that are tangent to the image surface f divide into two systems,—one, a system of double tangents, and the other, the assemblage of lines which are tangent to f in the points of a certain curve. But this curve, being the image of an imaginary line surface which touches F along a curve of curvature, is one of the principal tangent curves of f.

This determination of a principal tangent curve becomes illusory, however, if the focal plane—or, more correctly, a reducible part of it—, and not the congruence, is given arbitrarily. For on

[1] *Comptes rendus* (1864).

[2] Klein and Lie, in *Berliner Monatsbericht,* Dec. 15, 1870.

[3] Clebsch has determined the principal tangent curves of the Steiner surface.

a surface there is ordinarily only a finite number of points at which the tangent plane is also the plane associated, through a given linear complex, with that point.

It is of interest to note that a line surface whose genetrices belong to a linear complex contains a singly infinite set of points for each of which the tangent plane is also the plane associated, through the linear complex, with that point. The assemblage of these points forms a principal tangent curve, determinable by simple operations,—differentiation and elimination.

Now, Mr. Clebsch has demonstrated that if one principal tangent curve is known on a line surface, the others may be found by quadrature.

The determination of the principal tangent curves on a line surface belonging to a linear complex depends only on quadrature.

Applying our principle of transformation to the statement quoted from Clebsch as well as to its corollary proposition we arrive at the following theorems:

If on a tubular surface (sphere envelope) a non-circular curve of curvature is known, the others may be found by quadrature.

A singly infinite set of spheres which intersect a given sphere S at a constant angle envelope a tubular surface on which one curve of curvature can be given and the others obtained by quadrature.

That a curve of curvature can be found on the tubular surface is apparent also from the fact that the tubular surface intersects S at a constant angle. This curve of intersection must be one of the curves of curvature of the tubular surface, according to a certain well-known proposition. This proposition states: If two surfaces intersect at a constant angle, and the intersection curve is a line of curvature on one surface it is also such a line on the other. But on a sphere every curve is a line of curvature.

§12

The Correspondence between the Transformations of the Two Spaces

30. We may, as stated in Article 16, express our transformations by means of five equations which in the two groups (x, y, z, p, q) (X, Y, Z, P, Q) determine any quantity in one as a function of quantities in the other. If one of the two spaces, r, for example, undergoes a transformation, in which surfaces that are tangent are changed into similar surfaces, the corresponding transformation

of the other space will possess the same property. For, the transformation of r may be expressed by five equations in x_1, y_1, z_1, p_1, q_1, and x_2, y_2, z_2, p_2, q_2,—the subscripts 1 and 2 refer to the two states of r—and these relations are changed by the aid of the representation equation in x, y, z, p, q and X, Y, Z, P, Q, into relations in X_1, Y_1, Z_1, P_1, Q_1 and X_2, Y_2, Z_2, P_2, Q_2. Ths proves our assertion.

If we limit ourselves to linear transformations of r, we find among the corresponding transformations of R the following: *all movements (translational, rotational, and helicoidal), the transformation of similarity, the transformation by reciprocal radii, the parallel transformation*[1] (*transferring from one surface to a parallel surface*), *a reciprocal transformation studied by Mr. Bonnet.*[2] All of these, since they correspond to linear transformations in r, possess the property that they change curves of curvature into curves of curvature. We shall now prove *that to the general linear transformation of r there corresponds the most general transformation of R in which lines of curvature are covariant curves.*

31. In the first place, consider the linear point transformations of r to which correspond linear point transformations of R. It is clear that here we meet only with those transformations of R in which the infinitely distant imaginary circle remains unchanged; but these we do obtain.

For such a linear point transformation of R carries, on the one hand, right lines intersecting the circle into similar right lines, and, on the other hand, spheres into spheres. Thus the corresponding transformation of r is at one and the same time both a point and a line transformation,—that is, a linear point transformation. Which was to be proved.

The general linear transformation of R which does not displace the infinitely distant circle includes seven constants; and it can be built up by translations and rotations in conjunction with the similarity transformation. The corresponding transformation of r, which obviously also involves seven constants, may be characterized by saying that it carries a linear complex $H = 0$ and a certain one of its lines (the infinitely distant line of the xy-plane) into itself. We could also define this transformation by saying that it carries a special linear congruence into itself.

---

[1] Bonnet's "dilation."

[2] *Comptes rendus.* Several times in the 1850's.

The linear point transformation of r corresponding to a translation of R may be determined analytically. A translation is expressed by these equations:

$$X_1 = X_2 + A; \quad Y_1 = Y_2 + B; \quad Z_1 = Z_2 + C; \quad H_1 = H_2.$$

These equations and formulas (17) give the relations

$$r_1 = r_2 + a; \quad s_1 = s_2 + b; \quad \rho_1 = \rho_2 + c; \quad \sigma_1 = \sigma_2 + d.$$

Substituting these expressions in the equations of a right line,

$$r_1 z_1 = x_1 - \rho_1, \quad s_1 z_1 = y_1 - \sigma_1,$$

we obtain, as the definition of the required transformation of r, the following:

$$z_1 = z_2; \quad x_1 = x_2 + az_2 + c; \quad y_1 = y_2 + bz_2 + d.$$

It is likewise an easy matter to determine analytically the transformation of r corresponding to a *similarity transformation* of R. For, by applying (17), the equations

$$X_1 = mX_2; \quad Y_1 = mY_2; \quad Z_1 = mZ_2; \quad H_1 = mH_2$$

give the relations

$$r_1 = mr_2; \quad \rho_1 = m\rho_2; \quad s_1 = ms_2; \quad \sigma_1 = m\sigma_2.$$

These relations define a linear transformation of r which may also be expressed by the equations

$$z_1 = z_2; \quad x_1 = mx_2; \quad y_1 = my_2.$$

But these last equations define a linear point transformation which may be characterized by saying that in it *the points of two right lines remain stationary.*

By geometric considerations we shall show that rotations of R are also changed into transformations of the kind just described. Let A be the axis of rotation and M and N the two points of the imaginary circle that are not displaced by the rotation. It is clear that all the imaginary lines which intersect A and pass through M and N keep their position during the rotation. It follows that the same obtains for the image points of these lines, which form two right lines parallel to the xy-plane.

32. Transformation of the space R by reciprocal radii carries points into points, spheres into spheres and, finally, right lines of length zero into similar lines. The corresponding transformation of r is therefore a *linear point* transformation which carries the complex $H = 0$ into itself. If we note further that in the

transformation by reciprocal radii the points and right line generatrices of a certain sphere keep their position, it is clear that the corresponding reciprocal point transformation will not displace the points of the two right lines.

Mr. Klein[1] has called attention to the fact that the transformation just mentioned may be thought of as consisting of two transformations relative to two linear complexes in involution. In this case, $H = 0$ is one complex; the other is the one that corresponds to the assemblage of spheres which intersect orthogonally the fundamental sphere of the given reciprocal radii transformation.

From which it is clear that to a surface F which is carried into itself by a reciprocal radii transformation, there corresponds in space r a congruence belonging to $H = 0$, which is its own reciprocal polar in respect to a linear complex in involution with $H = 0$. The focal surface (f) of the congruence in question is thus its own reciprocal polar in respect to both the linear complexes. Consequently the totality of the double tangents of f is generally broken up into three congruences, two of which belong to $H = 0$ and to the complex in involution with $H = 0$.

33. Now consider, on the one hand, all line transformations of r by which right lines that intersect one another are changed into similar lines[2] and, on the other, the corresponding transformations of R which possess the property that they change spheres into spheres and spheres that are tangent into similarly places spheres.

This line transformation changes the assemblage of tangents to a surface f_1 into the totality of tangents to another surface f_2. Particularly, the principal tangents of f_1 change into the principal tangents of f_2,—this irrespective of whether the line transformation considered is a point transformation or a point-plane transformation.

By the corresponding transformation of R the triple infinity of spheres that touch a surface F_1 is changed into the totality of spheres which have a similar relation to F_2; and, specifically, the principal spheres of F_2. From this it follows that there is a correspondence between the lines of curvature of F_1 and F_2, in the sense that if in a relation $\phi(X_1, Y_1, Z_1, P_1, Q_1) = 0$, valid along a line of curvature for F_1, we substitute for X_1, Y_1, Z_1, P_1, Q_1 the values

---

[1] "Zur Theorie ———," in *Math. Annalen*, Vol. II.

[2] We must here consider two essentially different cases; for lines that are concurrent may be changed either into similarly placed lines or into lines that are coplanar.

X_2, Y_2, Z_2, P_2, Q_2, we obtain an equation which is valid for one of the curves of curvature of F_2.

I shall now prove that every transformation of R of the form

$$X_1 = F_1\left(X_2, Y_2, Z_2, \frac{dZ_2}{dX_2}, \frac{dZ_2}{dY_2}, \frac{d^2Z_2}{dX_2} \cdots \frac{d^{m+n}Z_2}{dX_2{}^m dY_2{}^n}\right)$$
$$Y_1 = F_2\left(X_2, Y_2, Z_2, \ldots\ldots\ldots\ldots\ldots \frac{d^{m+n}Z_2}{dX_2{}^m dY_2{}^n}\right)$$
$$Z_1 = F_3\left(X_2, Y_2, Z_2, \ldots\ldots\ldots\ldots\ldots \frac{d^{m+n}Z_2}{dX_2{}^m dY_2{}^n}\right)$$

which changes the lines of curvature of any given surface into lines of curvature of the new surface, corresponds, by my representation, to a linear transformation of r.

The proof of this reduces at once to showing that if a transformation of r changes the principal tangent curves of any surface into principal tangent curves of the transformed surface, then intersecting right lines are changed by the same transformation into similarly situated lines.

To begin with, the transformation in question must change right lines into right lines; because the right line is the only curve which is a principal tangent curve for every surface containing same.

Furthermore, that to right lines that intersect correspond right lines in the same relative position may be deduced from the fact that the developable surface is the only line surface so constituted that through each of its points passes only one principal tangent curve. Our transformation, therefore, changes developable surfaces into developable surfaces.

Hence, our statement is proved.

It may be remarked that, corresponding to the two essentially different kinds of linear *transformations there exist two distinct classes of transformations for which curves of curvature are covariant curves.*

If among the aforementioned transformations of R we choose those that are point transformations, we obtain *the most general point transformation of R in which lines of curvature are covariant curves*, a problem first solved by *Liouville.* That conformity is preserved even in the smallest parts is due to the fact that infinitesimal spheres carry into infinitesimal spheres.

The *parallel transformation* is known to carry lines of curvature into lines of curvature, and it is in reality easy to verify that the corresponding transformation of r is a linear point transformation.

For the equations

$$X_1 = X_2; \qquad Y_1 = Y_2; \qquad Z_1 = Z_2; \qquad H_1 = H_2 + A$$

are transformed (compare with our observations on translation in article 31) into relations of the form

$$z_1 = z_2; \qquad x_1 = x_2 + az_2 + b; \qquad y_1 = y_2 + cz_2 + d.$$

34. *Mr. Bonnet* has frequently discussed a transformation which he defines by the equations

$$Z_2 = iZ_2\sqrt{1 + p_2^2 + q_2^2}; \quad x_1 = x_2 + p_2z_2; \quad y_1 = y_2 + q_2z_2,$$

where the two subscripts refer to the given and to the transformed surface.

He proves that this transformation is a reciprocal one, in the sense that if applied twice it leads back to the given surface; that it transforms lines of curvature into lines of curvature; that, finally, if H_1 and H_2 indicate radii of curvature at corresponding points and if ζ_1 and ζ_2 are z-ordinates of the corresponding centers of curvature, these relations come about:

$$\zeta_1 = iH_2, H_1 = -i\zeta_2 \quad (\alpha)$$

Bonnet's transformation is the image of a transformation of r in respect to the linear complex $Z + iH = 0$. This we shall prove. If we recall that $X = 0$, $Y = 0$, $Z = 0$, $H = 0$ are in involution by pairs, we shall find that the coordinates of two right lines which are mutually polars in respect to $Z + iH = 0$ satisfy these relations:

$$X_1 = X_2; \qquad Y_1 = Y_2; \qquad Z_1 = iH_2; \qquad H_1 = -iZ_2. \quad (\beta)$$

But if X, Y, Z, H are interpreted as sphere coordinates, these formulas determine a relation by pairs among all the spheres of the space, precisely the same as the transformation of Bonnet.

For the principal spheres of a surface F_1 are by this changed into the principal spheres of surface F_2, and herein we recognize Bonnet's formulas (α). Moreover, if we think of F_1 as generated by point spheres, then the equations (β) define F_2 as an envelope of spheres whose centers lie in the plane $Z = 0$; for $Z_2 = 0$, since $H_1 = 0$. This leads exactly to the geometric construction given by Mr. Bonnet.

MÖBIUS, CAYLEY, CAUCHY, SYLVESTER, AND CLIFFORD

On Geometry of Four or More Dimensions

(Selections and translations made by Professor Henry P. Manning, Brown University, Providence, R. I.)

All references to a geometry of more than three dimensions before 1827 are in the form of single sentences pointing out that we cannot go beyond a certain point in some process because there is no space of more than three dimensions, or mentioning something that would be true if there were such a space. For the next fifty or sixty years the subject is treated more positively, but still in a fragmentary way, single features being developed to be used in some memoir on a different subject. The following selections are from some of the more interesting of those memoirs which of themselves, and because of the standing of the authors in the mathematical world, were to have apparently the chief influence in the further growth of this subject. The first article, by Möbius, is from *Der barycentrische Calcul* (Leipzig, 1827), a work from which other extracts have been made on pages 670–677 for another purpose. A brief biographical note accompanies that translation.

MÖBIUS

On Higher Space[1]

§139, page 181. If, given two figures, to each point of one corresponds a point of the other so that the distance between any two points of one is equal to the distance between the corresponding points of the other, then the figures are said to be *equal and similar.*

. . . .

§140, pages 182–183. *Problem.—To construct a system of n points which is equal and similar to a given system of n points.*

Solution. Let A, B, C, D,. . ., be the points of the given system, and A', B', C', D',. . ., the corresponding points of the system to be constructed. We have to distinguish three cases according as the points of the first set lie on a line, or in a plane, or in space.

. . . .

Finally, if the given system lies in space, then A' is entirely arbitrary, B' is an arbitrary point of the spherical surface which has A' for center and AB for radius, C' is an arbitrary point of the circle in which the two spherical surfaces drawn from A' with AC as radius and from B' with BC as radius intersect, and D' is one of the two points in which the three spherical surfaces drawn from A' with AD, from B' with BD, and from C' with CD as radii intersect. In the same way as D' will also each of the remaining points, for example, E', be found, only that of the two common intersections of the spherical surfaces drawn from A', B', C', with AE, BE, CE as radii, that one is taken which lies on the same side or on the opposite side of the plane $A'B'C'$ as D', according as the one or the other is the case with the corresponding points in the given system.

For the determination of A' therefore no distance is required, for the determination of B' one, for the determination of C' two, and for the determination of each of the remaining $n - 3$ points three. Therefore in all

$$1 + 2 + 3(n - 3) = 3n - 6$$

distances are required.

Remark.—Thus only for the point D', and for none of the remaining points, are we free to choose between the two intersections

[1] (From *Der barycentrische Calcul,* Leipzig, 1827, part 2, Chapter 1.)

on the three spherical surfaces falling on opposite sides of the plane $A'B'C'$. These two intersections are distinguished from each other in this way, that looking from one the order of the points A', B', C' is from right to left, but from the other from left to right, or, as also we can express it, that the former point lies on the left, the latter on the right of the plane $A'B'C'$. Now according as we choose for D' the one or the other of these two points, so also will the order formed be the same or different from that in which the point D appears from the points A, B, C. In both cases are the systems $A, B, C, D, \ldots$, and $A', B', C', D', \ldots$ indeed equal and similar, but only in the first case can they be brought into coincidence.

It seems remarkable that solid figures can have equality and similarity without having coincidence, while always, on the contrary, with figures in a plane or systems of points on a line equality and similarity are bound with coincidence. The reason may be looked for in this, that beyond the solid space of three dimensions there is no other, none of four dimensions. If there were no solid space, but all space relations were contained in a single plane, then would it be even as little possible to bring into coincidence two equal and similar triangles in which corresponding vertices lie in opposite orders. Only in this way can we accomplish this, namely by letting one triangle make a half revolution around one of its sides or some other line in its plane, until it comes into the plane again. Then with it and the other triangle will the order of the corresponding vertices be the same, and it can be made to coincide with the other by a movement in the plane without any further assistance from solid space.

The same is true of two systems of points $A, B, \ldots$ and $A', B', \ldots$ on one and the same straight line. If the directions of AB and $A'B'$ are opposite, then in no way can a coincidence of corresponding points be brought about by a movement of one system along the line, but only through a half revolution of one system in a plane going through the line.

For the coincidence of two equal and similar systems, $A, B, C, D, \ldots$ and $A', B', C', D', \ldots$ in space of three dimensions, in which the points $D, E, \ldots$ and $D', E', \ldots$ lie on opposite sides of the planes ABC and $A'B'C'$, it will be necessary, we must conclude from analogy, that we should be able to let one system make a half revolution in a space of four dimensions. But since such a space cannot be thought, so is also coincidence in this case impossible.

CAYLEY

On Higher Space

Arthur Cayley (1821–1895) was Sadlerian professor of mathematics at Cambridge. He wrote memoirs on nearly all branches of mathematics and, in particular created the theory of invariants. The extract is from his memoir, written in French, "On some Theorems of Geometry of Position," Crelle's *Journal*, vol. 31, 1846, pp. 213–227; *Mathematical Papers*, vol. I, Number 50, pp. 317–328.

In taking for what is given any system of points and lines we can draw through pairs of given points new lines, or find new points, namely, the points of intersection of pairs of given lines, and so on. We obtain in this way a new system of points and lines, which can have the property that several of the points are situated on the same line or several of the lines pass through the same point, which gives rise to so many theorems of the geometry of position. We have already studied the theory of several of these systems; for example, that of four points, of six points situated by twos on three lines which meet in a point, of six points three by three on two lines, or, more generally, of six points on a conic (this last case that of the mystic hexagram of Pascal, is not yet exhausted, we shall return to it in what follows), and also some systems in space. However, there exist systems more general than those which have been examined, and whose properties can be perceived in a manner almost intuitive, and which, I believe, are new.

Commence with the case most simple. Imagine a number n of points situated in any manner in space, which we will designate by 1, 2, 3, . . . n. Let us pass lines through all the combinations of two points, and planes through all the combinations of three points. Then cut these lines and planes by any plane, the lines in points and the planes in lines. Let $\alpha\beta$ be the point which corresponds to the line drawn through the two points α and β, let $\beta\gamma$ be the point which corresponds to that drawn through β and γ and so on. Further let $\alpha\beta\gamma$ be the line which corresponds to the plane passed through the three points α, β, and γ, etc. It is clear that the three points $\alpha\beta$, $\alpha\gamma$, and $\beta\gamma$ will be situated on the line

$\alpha\beta\gamma$. Then, representing by N_2, N_3, ... the numbers of the combinations of n letters 2 at a time, 3 at a time, etc., we have the following theorem.

THEOREM I.—*We can form a system of N_2 points situated* 3 *at a time on N_3 lines, to wit, representing the points by* 12, 13, 23, *etc., and the lines by* 123, *etc., the points* 12, 13, 23, *will be situated on the line* 123, *and so on.*

For $n = 3$ and $n = 4$ this is all very simple; we have three points on a line, or 6 points, 3 at a time, on 4 lines. There results no geometrical property. For $n = 5$ we have 10 points, 3 at a time on as many lines, to wit the points

12 13 14 15 23 24 25 34 35 45,

and the lines

123 124 125 134 135 145 234 235 245 345.

The points 12, 13, 14, 23, 24, 34 are the angles of an arbitrary quadrilateral,[1] the point 15 is entirely arbitrary, the point 25 is situated on the line passing through the points 12 and 15, but its position on this line is arbitrary. We will determine then the points 35 and 45, 35 as the point of intersection of the line passing through 13 and 15 and the line passing through 23 and 25, that is, of the lines 135 and 235, and the point 45 as the point of intersection of the lines 145 and 245. The points 35 and 45 will have the geometrical property of being in a line with 34, or all three will be in the same line 345.

. . . .

Page 217.

The general theorem, Theorem I, can be considered as the expression of an analytical fact, which ought equally well to hold in considering four coordinates instead of three. Here a geometrical interpretation holds which is applied to the points in space. We can, in fact, without having recourse to any metaphysical notion in regard to the possibility of a space of four dimensions, reason as follows (all of this can also be translated into language purely analytical): In supposing four dimensions of space it is necessary to consider *lines* determined by two points, *half-planes* determined by three points, and *planes*[2] determined by four points

[1] It is necessary always to have regard to the difference between quadrilateral and quadrangle. Each quadrilateral has four sides and six angles; each quadrangle has four angles and six sides.

[2] [His plane is what we call a hyperplane and his half-plane is an ordinary plane, and so he has to distinguish between a plane and an ordinary plane.]

(two planes intersect in a half-plane, etc.). Ordinary space can be considered as a plane, and it will cut a plane in an ordinary plane, a half-plane in an ordinary line and a line in an ordinary point. All this being granted, let us consider a number, n, of points, combining them by two, three, and four, in lines, half-planes, and planes, and then cut the system by space considered as a plane. We obtain the following theorem of geometry of three dimensions:

THEOREM VII.—*We can form a system of* N_2 *points, situated* 3 *by* 3 *in* N_3 *lines which themselves are situated* 4 *by* 4 *in* N_4 *planes. Representing the points by* 12, 13, *etc., the points situated in the same line are* 12, 13, 23, *and lines being represented by* 123, *etc., as before, the lines* 123, 124, 134, 234 *are situated in the same plane,* 1234.

In cutting this figure by a plane we obtain the following theorem of plane geometry:

THEOREM VIII.—*We can form a system of* N_3 *points situated* 4 *by* 4 *on* N_4 *lines. The points ought to be represented by the notation* 123, *etc., and the lines by* 1234, *etc. Then* 123, 124, 134, 234 *are in the same line designated by* 1234.

. . . .

CAUCHY

On Higher Space

When Louis Phillippe came to the throne Augustin Louis Cauchy (1789–1857), was unwilling to take the oath required by the government and for a while was in exile in Switzerland and Italy, but he returned to Paris in 1838 and finally became professor at the École Polytechnique. For a further biographical note see page 635. The article here translated is his "Memoir on Analytic Loci, *Comptes Rendus*, vol. XXIV, p. 885 (May 24, 1847); *Complete Works*, first series, vol. X, Paris, 1897, p. 292. It is one of a collection of memoirs on radical polynomials, a radical polynomial being a polynomial

$$\alpha + \beta\rho + \gamma\rho^2 + \ldots \eta\rho^{n-1},$$

where ρ is a primitive root of the equation

$$x^n = 1.$$

Consider several variables, $x, y, z, \ldots$ and various explicit functions, $u, v, w, \ldots$ of these variables. To each system of values of the variables $x, y, z, \ldots$ will generally correspond determined values of the functions $u, v, w, \ldots$ Moreover, if the variables are in number only two or three they can be thought of as representing the rectangular coordinates of a point situated in a plane or in space, and therefore each system of values of the variables can be thought of as corresponding to a determined point. Finally, if the variables x, y or x, y, z are subject to certain conditions represented by certain inequalities, the different systems of values of x, y, z for which the conditions are satisfied will correspond to different points of a certain locus, and the lines or surfaces which limit this locus in the plane in question or in space will be represented by the equations into which the given inequalities are transformed when in them we replace the sign $<$ or $>$ by the sign $=$.

Conceive now that the number of variables $x, y, z, \ldots$ becomes greater than three. Then each system of values of $x, y, z, \ldots$ will determine what we shall call an *analytical point* of which these variables are the coordinates, and to this point will correspond a certain value of each function of $x, y, z, \ldots$ Further, if the variables are subject to conditions represented by inequalities, the systems of values of $x, y, z, \ldots$ for which these conditions are satisfied will

correspond to analytical points, which together will form what we shall call an *analytical locus*. Moreover, this locus will be limited by analytical envelopes whose equations will be those to which the given inequalities are reduced when in them we replace the sign $<$ or $>$ by the sign $=$.

We shall also call *analytical line* a system of analytical points whose coordinates are expressed by aid of given linear functions of one of them. Finally, the *distance* of two analytical points will be the square root of the sum of the squares of the differences between the corresponding coordinates of these two points.

The consideration of analytical points and loci furnishes the means of clearing up a great many delicate questions, and especially those which refer to the theory of radical polynomials.

. . . .

SYLVESTER

On Higher Space

James Joseph Sylvester (1814–1897) was barred from certain honors in England because he was a Jew. He was professor at the University College, London, and at the Royal Military Academy in Woolwich. For a short time he taught at the University of Virginia. When the Johns Hopkins University was started he went there to take the lead in the advance of higher mathematics in this country. In 1883 he returned to England and became Savilian professor of geometry at Oxford. The article quoted is "On the Center of Gravity of a Truncated Triangular pyramid, and on the Principles of Barycentric Perspective." It appeared in the *Philosophical Magazine*, vol. XXVI, 1863, pp. 167–183; *Collected Mathematical Papers*, vol. II, Cambridge, 1908, pp. 342–357.

There is a well-known geometrical construction for finding the center of gravity of a plane quadrilateral which may be described as follows.

Let the intersection of the two diagonals (say Q) be called the *cross-center*, and the intersection of the lines bisecting opposite sides (say O) the *mid-center* (which, it may be observed, is the center of gravity of the four angles viewed as equal weights), then the center of gravity is in the line joining these two centers produced past the latter (the mid-center), and at a distance from it equal to one-third of the distance between the two centers. In a word, if G be the center of gravity of the quadrilateral, QOG will be in a right line and $OG = \frac{1}{3}QO$.

The frustum of a pyramid is the nearest analogue in space to a quadrilateral in a plane since the latter may be regarded as the frustum of a triangle. The analogy, however, is not perfect, inasmuch as a quadrilateral may be regarded as a frustum of either of two triangles, but the pyramid to which a given frustum belongs is determinate. Hence *à priori* reasonable doubts might have been entertained as to the possibility of extending to the pyramidal frustum the geometrical method of centering the plane quadrilateral. The investigation subjoined dispels this doubt, and will be found to lead to the perfect satisfaction, under a somewhat unexpected form, of the hoped-for analogy.

Let abc and $\alpha\beta\gamma$ be the two triangular faces, and $a\alpha$, $b\beta$, and $c\gamma$ the edges of the quadrilateral faces of a pyramidal frustum. Then this frustum may be resolved in six different ways into three different pyramids as shown in the annexed double triad of schemes.

$$\begin{array}{cccc} a & b & c & \alpha \\ b & c & \alpha & \beta \\ c & \alpha & \beta & \gamma \end{array} \qquad \begin{array}{cccc} b & c & a & \beta \\ c & a & \beta & \gamma \\ a & \beta & \gamma & \alpha \end{array} \qquad \begin{array}{cccc} c & a & b & \gamma \\ a & b & \gamma & \alpha \\ b & \gamma & \alpha & \beta \end{array}$$

$$\begin{array}{cccc} b & a & c & \beta \\ a & c & \beta & \alpha \\ c & \beta & \alpha & \gamma \end{array} \qquad \begin{array}{cccc} a & c & b & \alpha \\ c & b & \alpha & \gamma \\ b & \alpha & \gamma & \beta \end{array} \qquad \begin{array}{cccc} c & b & a & \gamma \\ b & a & \gamma & \beta \\ a & \gamma & \beta & \alpha \end{array}$$

If then, taking any one of the above schemes, we draw a plane through the centers[1] of the three pyramids of which it is composed, the six planes thus drawn will meet in a point, which will be the center of the frustum.[2]

Let the point in which αa, βb, and γc meet when produced be the origin of coordinates, and $bc\beta\gamma$, $ca\gamma\alpha$, and $ab\alpha\beta$ be taken as the planes of x, y, z and let $4a, 0, 0; 0, 4b, 0; 0, 0, 4c$ be the coordinates of a, b, c, and $4\alpha, 0, 0; 0, 4\beta, 0; 0, 0, 4\gamma$ those of α, β, γ. Consider the first of the schemes above written.

$a + \alpha, b, c$ will be the coordinates of the center of $abc\alpha$,
$\alpha, b + \beta, c$ will be the coordinates of the center of $bc\alpha\beta$,
$\alpha, \beta, c + \gamma$ will be the coordinates of the center of $c\alpha\beta\gamma$,

because, as everyone knows, the center of a triangular pyramid is the same as that of its angles regarded as of equal weight. But again, if we define as the *mid-center* the center of the six angles of the frustum regarded as of equal weight, its coordinates will be

$$\frac{2a + 2\alpha}{3}, \frac{2b + 2\beta}{3}, \frac{2c + 2\gamma}{3};$$

and if we substitute for each of the three centers last named, points lying, respectively, in a right line with them and the mid-center, on the opposite side of the mid-center and at distances from it double

[1] I shall throughout in future for greater brevity hold myself at liberty to use the word center to mean center of gravity.

[2] I shall hereafter show that these six planes all touch the same cone, of which, as also of its polar reciprocal, I have succeeded in obtaining the equation.

those of these centers themselves, these quasi-images of the centers in question will have for their coordinates

$$\begin{gathered} 0,\ 2\beta,\ 2\gamma, \\ 2a,\ 0,\ 2\gamma, \\ 2a,\ 2b,\ 0. \end{gathered}$$

These points are, accordingly, the centers of the lines $\beta\gamma$, γa, and ab, respectively.

And a similar conclusion will apply to each of the six schemes. Hence, using in general (p, q) to mean the middle of the line pq, and by the collocation of the symbols for three points understanding the plane passing through them, it is clear

1. That the six planes

$$\begin{array}{lll} (\beta, \gamma)\ (\gamma, a)\ (a, b) & (\gamma, \alpha)\ (\alpha, b)\ (b, c) & (\alpha, \beta)\ (\beta, c)\ (c, a) \\ (\gamma, \beta)\ (\beta, a)\ (a, c) & (\alpha, \gamma)\ (\gamma, b)\ (b, a) & (\beta, \alpha)\ (\alpha, c)\ (c, b) \end{array}$$

will meet in a single point which may be called the *cross-center*, being the true analogue of the intersection of the two diagonals of a quadrilateral figure in the plane.

2. That if we join this cross-center (say Q) with O the mid-center, and produce QO to G, making $OG = \frac{1}{2}QO$, G will be the center of the frustum $abc\alpha\beta\gamma$.[1]

It may be satisfactory to some of my readers to have a direct verification of the above.

Let then

$$A = \frac{a^2bc - \alpha^2\beta\gamma}{abc - \alpha\beta\gamma},\ B = \frac{ab^2c - \alpha\beta^2\gamma}{abc - \alpha\beta\gamma},\ C = \frac{abc^2 - \alpha\beta\gamma^2}{abc - \alpha\beta\gamma}.$$

A moment's reflection will serve to show that A, B, C are the coordinates of the center of the frustum.[2]

---

[1] [The three centers of the three tetrahedrons lie in a plane through the center G. Drawing lines from these three points and G to the mid-center O, and laying off on these lines produced beyond O any given multiples of these lines, we shall have three points corresponding to the three given points and a fourth point Q corresponding to G, all lying in a plane parallel to the first plane. We can do this for any three of the six planes through G and get planes whose intersection will be Q, and then from Q and O we can get G by reversing the process. If we take the multiplier to be 2, the three new points will be very simple as pointed out in the footnote.]

[2] [These expressions can be obtained, for example, by considering the frustum as the difference of two pyramids with a common vertex at the origin.]

Again, the first three of the six planes last referred to will be found to have for their equations, respectively,

$$\begin{aligned} \beta\gamma x + \gamma a y + a b z &= 2a\gamma(b + \beta), \\ bcx + \gamma\alpha y + \alpha b z &= 2b\alpha(c + \gamma), \\ \beta c x + c a y + \alpha\beta z &= 2c\beta(a + \alpha). \end{aligned}$$

The determinant

$$\begin{vmatrix} \beta\gamma & \gamma a & ab \\ bc & \gamma\alpha & \alpha b \\ \beta c & ca & \alpha\beta \end{vmatrix} = (abc - \alpha\beta\gamma)^2.$$

The determinant

$$\begin{vmatrix} \gamma a & ab & 2a\gamma(b + \beta) \\ \gamma\alpha & \alpha b & 2b\alpha(x + \gamma) \\ ca & \alpha\beta & 2c\beta(a + \alpha) \end{vmatrix}$$

$$= 2\alpha a(bc - \beta\gamma)(abc - \alpha\beta\gamma) = 2[(\alpha^2\beta\gamma - a^2bc)(abc - \alpha\beta\gamma) + (a + \alpha)(abc - \alpha\beta\gamma)^2].$$

Hence if x, y, and z be the coordinates of the intersection of the above-mentioned three planes,

$$\begin{aligned} x &= -2A + 2(a + \alpha), \\ y &= -2B + 2(b + \beta), \\ z &= -2C + 2(c + \gamma), \end{aligned}$$

and the same will evidently be true of the other ternary system of planes, so that all six planes intersect in a single point Q, of which x, y, and z above written are the coordinates. And the coordinates of O being

$$\frac{2a + 2\alpha}{3}, \frac{2b + 2\beta}{3}, \frac{2c + 2\gamma}{3},$$

and those of G being A, B, C, it is obvious that QOG is a right line, and $OG = \frac{1}{2}QO$, as was to be shown.

The analogy with the quadrilateral does not end here. There is a construction[1] for the center of a quadrilateral still easier than

[1] This is the mode of statement (except that the important notion of opposite points was not explicitly contained in it) which, accidently meeting my eye in a proof sheet of some geometrical notes (by an anonymous author) intended for further insertion in the forthcoming (if not forthcome) number of the *Quarterly Journal of Mathematics*, led to the long train of reflections embodied in this paper, which but for that casual glance would never have seen the light. The same construction, under another and somewhat less eligible form, is given in the *Mathematician* (a periodical now extinct, edited by Dr. Rutherford and Mr. Fenwick, both of the Royal Military Academy), 1847, volume II, page 292, and is therein stated by the latter gentleman to have, "as he believes, first appeared in the *Mechanics Magazine*, and subsequently in the *Lady's Diary* for 1830."

that above cited, which may be expressed in general terms by aid of a simple definition. Agree to understand by the *opposite* to a point L on a limited line AB a point M such that L and M are at equal distances from the center of AB but on opposite sides of it. Then we may affirm that the center of a quadrilateral is the center of the triangle whose apices are the intersection of its two diagonals (that is, the cross-center) and the opposites of that intersection on these two diagonals, respectively. So now, if we agree to understand by opposite points on a limited triangle two points on a line with the center of the triangle and at equal distances from it on opposite sides, and bear in mind that the cross-center of a pyramidal frustum is the intersection of either of two distinct ternary systems of triangles which may be called the two systems of cross-triangles,[1] we may affirm that the center of a pyramidal frustum is the center of a pyramid whose apices are its cross-center, and the opposites of that center on the three components of either of its systems of cross-planes. This is easily seen, for if we take the first of the two systems, their respective centers will be

$$\frac{4a}{3},\ \frac{2b+2\beta}{3},\ \frac{4\gamma}{3};\ \frac{4\alpha}{3},\ \frac{4b}{3},\ \frac{2c+2\gamma}{3};\ \frac{2a+2\alpha}{3},\ \frac{4\beta}{3},\ \frac{4c}{3}.$$

Thus the three opposites to the cross-center whose coordinates are

$$-2A+2(a+\alpha),\ -2B+2(b+\beta),\ -2C+2(c+\gamma),$$

will have for their x coordinates[2]

$$\frac{2a}{3}-2\alpha+2A;\ -2a+\frac{2\alpha}{3}+2A;\ -\frac{2a}{3}-\frac{2\alpha}{3}+2A;$$

for their y coordinates

$$\frac{2b}{3}-2\beta+2B;\ -2b+\frac{2\beta}{3}+2B;\ -\frac{2b}{3}-\frac{2\beta}{3}+2B;$$

and for their z coordinates

$$\frac{2c}{3}-2\gamma+2C;\ -2c+\frac{2\gamma}{3}+2C;\ -\frac{2c}{3}-\frac{2\gamma}{3}+2C;$$

---

[1] From the description given previously, it will be seen that a cross-triangle of the frustum is one which has its apices at the centers of either diagonal of any quadrilateral face and of the two edges conterminous but not in the same face with that diagonal.

[2] These are not arranged so that the three coordinates of a point are in a column. There is a certain cyclical shifting in the second and third lines. If we think of the nine coordinates in the arrangement here as forming a determinant, we get the coordinates of the three opposites separately by taking the three negative diagonals.

and consequently the center of the pyramid whose apices are the cross-center and its three opposites will be A, B, C, that is, will be the center of gravity of the frustum, as was to be shown.[1]

It is clear that these results may be extended to space of higher dimensions. Thus in the corresponding figure in space of four dimensions bounded by the hyperplanar quadrilaterals[2] *abcd* and $\alpha\beta\gamma\delta$, which will admit of being divided into four hyperpyramids in 24 different ways, all corresponding to the type

$$\begin{matrix} a & b & c & d & \alpha \\ b & c & d & \alpha & \beta \\ c & d & \alpha & \beta & \gamma \\ d & \alpha & \beta & \gamma & \delta, \end{matrix}$$

---

[1] I at one time supposed that a, b, c, α, β, γ, formed two systems of diagonal planes, and that there were thus two cross-centers, and dreamed a dream of the construction for the center of gravity of the pyramidal frustum based upon this analogy, inserted (it is true as a conjecture only) in the *Quarterly Journal of Mathematics*, but the nature of things is ever more wonderful than the imagination of men's minds, and her secrets may be won, but cannot be snatched from her. Who could have imagined *à priori* that for the purposes of this theory a diagonal of a quadrilateral was to be viewed as a line drawn through two opposite angles of the figure regarded, not as themselves, but as their own center of gravity. Some of my readers may remember a single case of a similar autometamorphism which occurred to myself in an algebraical inquiry, in which I was enabled to construct the canonical form of a six-degreed binary quantic from an analogy based on the same for a four-degreed one, by considering the square of a certain function which occurs in the known form as consisting of two factors, one the function itself, the other a function morphologically derived from, but happening for that particular case to coincide with the function. The parallelism is rendered more striking from the fact of 4 and 6 being the numbers concerned in each system of analogies, those numbers referring to degrees in one theory and to angular points in the other. It is far from improbable that they have their origin in some common principle, and that so in like manner the parallelism will be found to extend in general to any quantic of degree $2n$, and the corresponding barycentric theory of the figure with $2n$ apices (n of them in one hyperplane and n in another), which is the problem of a hyperpyramid in space of n dimensions. The probability of this being so is heightened by the fact of the barycentric theory admitting, as is hereafter shown, of a descriptive generalization, descriptive properties being (as is well known) in the closest connection with the theory of invariants. Much remains to be done in fixing the canonic forms of the higher even-degreed quantics, and this part of their theory may hereafter be found to draw important suggestions from the hypergeometry above referred to, if the supposed alliance have a foundation in fact.

[2] [This word should be *tetrahedrons*.]

there will be a *cross-center* given by the intersection of any four out of 24 hyperplanes resoluble into six sets of four each,[1]—one such set of four being given in the scheme subjoined, where in general pqr means the point which is the center of (p, q, r) and the collocation of four points means the hyperplane passing through them, namely,

$$\begin{array}{l} \beta\gamma\delta\ \gamma\delta a\ \delta ab\ abc,^{2} \\ \gamma\delta\alpha\ \delta\alpha b\ \alpha bc\ bca, \\ \delta\alpha\beta\ \alpha\beta c\ \beta cd\ cdb, \\ \alpha\beta\gamma\ \beta\gamma d\ \gamma da\ dac. \end{array}$$

The *mid-center* will mean the center of the eight angles a, b, c, d, α, β, γ, δ, regarded as of equal weight, and to find the center of the hyper-pyramidal frustum we may either produce the line joining the cross-center with the mid-center through the latter and measure off three-fifths of the distance of the joining line on the part produced (as in the preceding cases we measured off two-fourths and one-third of the analogous distance) or we may take the four opposites of the cross-center on the four components of any one of the six systems of hyperplanar tetrahedrons of which it is the intersection, and find the center of the hyperpyramid so formed. The point determined by either construction will be the center of gravity of the hyperpyramidal frustum in question. And so for space of any number of dimensions. It will of course be seen that a general theorem of determinants[3] is contained in

---

[1] [The second, third, and fourth of a set may be obtained from the first by taking the cylical permutations of the Roman letters with the same permutations of the Greek letters.]

[2] [The last letters of these four lines should be c, d, a, b.]

[3] We learn indirectly from this how to represent under the form of determinants of the ith order, and that in a certain number of ways, the general expressions

$$(l_1l_2\ldots l_i - \lambda_1\lambda_2\ldots\lambda_i)^{i-1}$$

and

$$l_1\lambda_1(l_2l_3\ldots l_i - \lambda_2\lambda_3\ldots\lambda_i)(l_1l_2\ldots l_i - \lambda_1\lambda_2\ldots\lambda_i)^{i-2}$$

a strange conclusion to be able to draw incidentally from a hyper-theory of center of gravity! Thus, for example, on taking $i = 4$ we shall find

$$\begin{vmatrix} bcd & cd\alpha & d\alpha\beta & \alpha\beta\gamma \\ \beta\gamma\delta & cda & da\beta & a\beta\gamma \\ b\gamma\delta & \gamma\delta\alpha & dab & ab\gamma \\ bc\delta & c\delta\alpha & \delta\alpha\beta & abc \end{vmatrix} = (abcd - \alpha\beta\gamma\delta)^3.$$

the assertion that for space of n dimensions there will be $n!$ quasi-planes all intersecting in the same point, as also in the general relation connecting this point (the cross-center) with the mid-center and center of gravity, of each of which it is easy to assign the value of the coordinates in the general case.

. . . .

And again

$$\begin{vmatrix} \alpha d(bc + c\beta + \beta\gamma) & cd\alpha & d\alpha\beta & \alpha\beta\gamma \\ \beta a(cd + d\gamma + \gamma\delta) & cda & da\beta & a\beta\gamma \\ \gamma b(da + a\delta + \delta\alpha) & \gamma\delta\alpha & dab & ab\gamma \\ \delta c(ab + b\alpha + \alpha\beta) & c\delta\alpha & \delta\alpha\beta & abc \end{vmatrix} = a\alpha(bcd - \beta\gamma\delta)(abcd - \alpha\beta\gamma\delta)^2$$

The number of these representations will not be 24, that is $4!$, but only 12, the half of that number, because it will be easily seen that the cycle *abcd*, $\alpha\beta\gamma\delta$ will lead to the same determinants, only differently arranged, as the cycles *bcda*, $\beta\gamma\delta\alpha$. I believe the law is that the number of varieties of such representations is $i!$ or $\frac{1}{2}i!$ according as i is odd or even. The expression $ab - \alpha\beta$ at once conjures up the idea of a determinant. We now see that there is an equally natural determinantive representation, or system of representations, of $(abc - \alpha\beta\gamma)^2$, $(abcd - \alpha\beta\gamma\delta)^3$, etc.

CLIFFORD

On Higher Space

William Kingdon Clifford (1848–1879), was professor of mathematics and mechanics in University College, London from 1871 to the time of his death. The following article is his solution of a "Problem in Probability," in the *Educational Times*, January, 1866, Problem 1878, proposed by himself. *A line of length a is broken up into pieces at random; prove that* (1) *the chance that they cannot be made into a polygon of n sides is* $n2^{1-n}$; *and* (2) *the chance that the sum of the squares described on them does not exceed* $\frac{a^2}{(n-1)}$ *is*

$$\left(\frac{\pi}{n^2 - n}\right)^{\frac{1}{2}n-1} \frac{\Gamma(n)}{\Gamma\{\frac{1}{2}(n+1)\}}, \frac{1}{n^{\frac{1}{2}}}.$$

Solution by the proposer. November, 1866; reprinted in *Mathematical Questions with Solutions*, vol. VI, London, 1866, pp. 83–87; also in *Mathematical Papers*, London, 1882, pp. 601–607.

1. Let us define as follows. A point is taken *at random* on a (finite or infinite) straight line when the chance that the point lies on a finite portion of the line varies as the length of that portion. And a line is broken up at random when the points of division are taken at random.

Now the n pieces will always be capable of forming a polygon except when one of them is greater than the sum of all the rest, that is, greater than half the line. The first part of the question may therefore be stated thus: $n - 1$ *points are taken at random on a finite line; to find the chance that some one of the intervals shall be greater than half the line.*

. . . .

4. *Third Solution.*—To make this clear I will state first the previously known analogous solutions in the cases where $n = 3$ and $n = 4$. When the line is divided into three pieces, call them x, y, and z, and take their lengths for the coordinates of a point P in geometry of three dimensions. Then, since

$$x + y + z = a \tag{1}$$

and x, y, and z are all positive, the point P must be somewhere on the surface of the equilateral triangle determined on the plane (1) by the coordinate planes. Now consider those points on the

triangle for which $x > \frac{1}{2}a$. These are cut off by the plane $x = \frac{1}{2}a$, and it is easy to see that this plane cuts off from one corner of the triangle a similar triangle of half the linear dimensions, and therefore of one-fourth the area. Now there are three corners cut off. Their joint area is therefore three-fourths of the area of the triangle, and the chance required is accordingly $\frac{3}{4}$.

When the line is divided into four pieces, take the first three pieces as the coordinates of a point in space. Then we have $x + y + z < a$ and x, y, and z all positive. So the point must lie within the content of the tetrahedron bounded by the plane $x + y + z = a$ and the coordinate planes. Now if $x + y + z < \frac{1}{2}a$ the fourth piece must be greater than $\frac{1}{2}a$. The points for which this is the case are cut off by the plane $x + y + z = \frac{1}{2}a$ and it is easily seen as before that this plane cuts off from one corner of the tetrahedron a similar tetrahedron of half the linear dimensions, and therefore of one eighth the volume. So also the plane $x = \frac{1}{2}a$ cuts off from another corner a similar tetrahedron of half the linear dimensions. Since therefore there are four corners cut off, their joint volume is four-eighths or one-half of the volume of the tetrahedron, and the chance required is accordingly $\frac{1}{2}$.

5. Now consider the analogous case in geometry of n dimensions. Corresponding to a closed area and a closed volume we have something which I shall call a *confine*. Corresponding to a triangle and to a tetrahedron there is a confine with $n + 1$ corners or vertices which I shall call a *prime confine*[1] as being the simplest form of confine. A prime confine has also $n + 1$ *faces*, each of which is, not a plane, but a prime confine of $n - 1$ dimensions. Any two vertices may be joined by a straight line, which is an *edge* of the confine. Through each vertex pass n edges. A prime confine may be *regular*, which it is when any three vertices form an equilateral triangle; or *rectangular*, which it is when the edges through some one vertex are all equal and at right angles to one another.

To solve the question for general values of n we may adopt as a type either of the geometrical solutions given for the cases $n = 3$ and $n = 4$. First take the lengths of the n pieces for the coordinates of a point in geometry of n dimensions. Then, since their sum is a and they are all positive, the point must lie within a

[1] [The term now commonly used is *simplex*. In space of four dimensions this is a *pentahedroid*.]

certain regular prime confine of $n-1$ dimensions. The supposition that a certain piece is greater than $\frac{1}{2}a$ cuts off from one corner of the confine a similar confine of half the linear dimensions and therefore of 2^{1-n} times the content. As there are n corners, their joint content is $n2^{1-n}$ times the content of the confine. The chance required is consequently $n2^{1-n}$. Or take the lengths of the first $n-1$ pieces as the coordinates of a point in geometry of $n-1$ dimensions. The point will then lie within a certain rectangular confine of $n-1$ dimensions, and the investigation proceeds as before, the n corners being cut off in the same manner.

6. It will be seen that this third solution involves in a geometrical form the assumption of which some sort of proof was given in the first solution. Let us make this extension of our fundamental definition:—A point is taken at random in a (finite or infinite) space of n dimensions when the chance that the point lies in a finite portion of the space varies as the content of that portion. The assumption is that when the lengths of the pieces into which a line is broken up are taken as coordinates of a point, then if the line is broken up at random the point is taken at random and *vice-versa.* The proof of this assumption may be shown to involve a geometrical proposition equivalent to the integration by parts of the differential in Art. 3.[1]

Making this assumption, we may solve the second part of the question by the method of the third solution of the first part. I will first state the previously known analogous solution of the case where $n = 3$. The question in this case is, *If a line of length a be broken into three pieces at random find the chance that the sum of the squares of these pieces shall be less than* $\frac{1}{2}a^2$. Take the lengths of the three pieces for coordinates x, y, and z of a point P in geometry of three dimensions. Then, as before, the point must lie somewhere in the area of the equilateral triangle determined on the plane $x + y + z = a$ by the coordinate planes. But if also the sum of the squares of the pieces is less than a certain quantity m^2, then the point P must lie within a certain circle

[1] [The assumption of which "some sort of a proof was given in the first solution" is that the chance that the rth piece reckoning from one end of the line shall be greater than $\frac{1}{2}a$ is equal to the chance that the $(r+1)$th piece shall be greater than $\frac{1}{2}a$. In the second solution (Art. 3) the chance that the rth piece shall be greater than $\frac{1}{2}a$ is proved equal to the integral

$$\frac{(n-1)!}{(n-r)!(r-2)!}\int_0^{\frac{1}{2}a}\left(\frac{x}{a}\right)^{r-2}\left(\frac{1}{2}-\frac{x}{a}\right)^{n-r+1}\frac{dx}{a}.]$$

determined on the plane $x + y + z = a$ by the sphere $x^2 + y^2 + z^2 = m^2$. Now in the case where $m^2 = \frac{1}{2}a^2$ this circle is the circle inscribed in the equilateral triangle, so that the question reduces itself to this one:—

To find in terms of an equilateral triangle the area of its inscribed circle.

Now let us go a little further and consider the case in which $n = 4$. Here we shall have to take a point P in geometry of four dimensions. The point must lie somewhere in the regular tetrahedron determined on the hyperplane $x + y + z + w = a$ by the coordinate hyperplanes. If also the sum of the squares of the pieces is less than a certain quantity m^2, then the point P must lie within a certain sphere determined on the hyperplane $x + y + z + w = a$ by the quasi-sphere $x^2 + y^2 + z^2 + w^2 = m^2$. In the particular case where m is the perpendicular from the vertex on the base of a rectangular tetrahedron each of whose equal edges is of length a, or $m^2 = \frac{1}{3}a^2$ this sphere is the sphere inscribed in the regular tetrahedron.[1] The question is therefore reduced to this one:—

To find in terms of a regular tetrahedron the volume of its inscribed sphere.

Now a similar reduction holds in the general case; namely, the question can always be reduced to this one:—

To find in terms of a regular prime confine of $n - 1$ dimensions the content of its inscribed quasi-sphere.

This question I proceed to solve.

7. Let $n - 1 = p$. *The perpendicular from any vertex on the opposite face of a regular prime confine in p dimensions*

$$= \left(\frac{p+1}{2p}\right)^{\frac{1}{2}} \cdot \text{(edge)}.$$

For let O be the vertex in question, OA, OB, ... the p edges through O. Draw through each vertex A a space of $p - 1$ dimensions parallel to the face opposite to A. The p spaces thus drawn will intersect in a point P such that OP is the diagonal of a confine analogous to a parallelogram and to a parallelepiped. Then OP is p times the perpendicular from O on the opposite

[1] [Each face of the regular tetrahedron is the base of a rectangular tetrahedron formed with its vertices and the origin, and the perpendicular from the origin upon this face will be the radius of a hypersphere which is tangent to the face, and so intersects the hyperplane of the tetrahedron in the inscribed sphere.]

face of the regular confine; for the perpendicular is the projection of one edge at a certain angle, while OP is the projection at the same angle of a broken line consisting of p edges.[1]

We have also[2]

$$\begin{aligned} OP^2 &= OA^2 + OB^2 + OC^2 + \ldots + 2OA.OB \cos AOB + \ldots \\ &= \Sigma.OA^2 + \Sigma.OA.OB \text{ [since } \cos AOB = \tfrac{1}{2}, \text{ etc.]} \\ &= [p + \tfrac{1}{2}p(p-1)].OA^2 = \tfrac{1}{2}p(p+1).OA^2, \end{aligned}$$

therefore $(\text{perpendicular})^2 = \dfrac{OP^2}{p^2} = \dfrac{p+1}{2p}.(\text{edge})^2.$

If the confine were rectangular, or all the angles at O right angles, we should have $\cos AOB = 0$, etc., and so

$$(\text{perpendicular})^2 = \frac{1}{p}(\text{edge})^2 = \frac{a^2}{n-1},$$

which proves that the question does always reduce itself to the one now under consideration.

The content of a regular prime confine in p dimensions whose edge is a is[3]

$$= \frac{a^p}{p!}\left(\frac{p+1}{2^p}\right)^{\frac{1}{2}}$$

Suppose this formula true for $p - 1$ dimensions; that is, let

$$V_{p-1} = \frac{a^{p-1}}{(p-1)!}\left(\frac{p}{2^{p-1}}\right)^{\frac{1}{2}}.$$

Now, content of confine $= \dfrac{1}{p} \times$ perpendicular $\times$ content of face, or

$$V_p = \frac{a}{p}\left(\frac{p+1}{2p}\right)^{\frac{1}{2}}.V_{p-1} = \frac{a^p}{p!}\left(\frac{p+1}{2^p}\right)^{\frac{1}{2}}.$$

Hence the formula, if true for one value of p, is true for the next. It can be immediately verified in the case of $p = 1$. Therefore it is generally true.

The radius of the inscribed quasi-sphere $\rho = \dfrac{a}{\{2p(p+1)\}^{\frac{1}{2}}}.$

---

[1] [The perpendicular produced to p times its length will give us OP. In fact, if we take the edges through O for a system of oblique axes the coordinates of P will all be equal and the line OP must make equal angles with the axes.]

[2] [See Salmon's *Geometry of Three Dimensions*, fourth edition, Dublin, 1882, p. 11.]

[3] [The edge in our case is $a\sqrt{2}$, but the ratio of the inscribed quasisphere does not depend on the length of the edge.]

We can divide the regular confine into $p + 1$ equal confines, each having the center of the inscribed quasi-sphere for vertex, and the content of one of these $= \frac{\rho}{p} \times$ content of face. But the sum of them all is equal to the content of the whole confine. Hence $(p + 1)\rho$ = perpendicular of confine[1]

$$= a\left(\frac{p+1}{2p}\right)^{1/2}, \text{ or, } \rho = \frac{a}{\{2p(p+1)\}^{1/2}}$$

The content of the quasi-sphere $= \rho^p \dfrac{\{\Gamma(\frac{1}{2})\}^p}{\Gamma(\frac{1}{2}p + 1)}$.

For it is the value of $\iiint \ldots dx\, dy\, dz \ldots$, the integral being so taken as to give to the variables all values consistent with the condition that $x^2 + y^2 + z^2 + \ldots$ is not greater than ρ^2 (see Todhunter's *Integral Calculus*, Art.[2] 271). Let C_p denote this content. Then

$$C_p = \rho^p \frac{\{\Gamma(\frac{1}{2})\}^p}{\Gamma(\frac{1}{2}p + 1)} = \frac{a^p}{(2p^2 + 2p)^{\frac{1}{2}p}} \cdot \frac{\{\Gamma(\frac{1}{2})\}^p}{\Gamma(\frac{1}{2}p + 1)}.$$

Therefore

$$\frac{C_p}{V_p} = \left(\frac{\pi}{p^2 + p}\right)^{\frac{1}{2}p} \cdot \frac{\Gamma(p+1)}{\Gamma(\frac{1}{2}p + 1)} \cdot \frac{1}{(p+1)^{\frac{1}{2}}}.$$

Restore $n - 1$ for p and we get the answer to the question, namely,

$$\left(\frac{\pi}{n^2 - n}\right)^{\frac{1}{2}(n-1)} \cdot \frac{\Gamma(n)}{\Gamma\{\frac{1}{2}(n+1)\}} \cdot \frac{1}{n^{\frac{1}{2}}}.$$

8. The following are applications of the same method.

If a line be broken up at random into n pieces, the chance of an assigned two of them (the pth and qth from one end) being together greater than half the line is $n2^{1-n}$.

If n pieces be cut off at random, one from each of n equal lines, the chance that the pieces cannot be made into a polygon is $\dfrac{1}{(n-1)!}$.

---

[1] [Two confines having the same base are to each other as their altitudes.]

[2] [In some editions at least (4th and 7th) this is Art. 275.]

IV. FIELD OF PROBABILITY

Fermat and Pascal on Probability

(Translated from the French by Professor Vera Sanford, Western Reserve University, Cleveland, Ohio.)

Italian writers of the fifteenth and sixteenth centuries, notably Pacioli (1494), Tartaglia (1556), and Cardan (1545), had discussed the problem of the division of a stake between two players whose game was interrupted before its close. The problem was proposed to Pascal and Fermat, probably in 1654, by the Chevalier de Méré, a gambler who is said to have had unusual ability "even for the mathematics." The correspondence which ensued between Fermat and Pascal, was fundamental in the development of modern concepts of probability, and it is unfortunate that the introductory letter from Pascal to Fermat is no longer extant. The one here translated, written in 1654, appears in the *Œuvres de Fermat* (ed. Tannery and Henry, Vol. II, pp. 288–314, Paris, 1894) and serves to show the nature of the problem. For a biographical sketch of Fermat, see page 213; of Pascal, page 67. See also pages 165, 213, 214, and 326.

Monsieur,

If I undertake to make a point with a single die in eight throws, and if we agree after the money is put at stake, that I shall not cast the first throw, it is necessary by my theory that I take $1/6$ of the total sum to be impartial because of the aforesaid first throw.

And if we agree after that that I shall not play the second throw, I should, for my share, take the sixth of the remainder that is $5/36$ of the total.

If, after that, we agree that I shall not play the third throw, I should to recoup myself, take $1/6$ of the remainder which is $25/216$ of the total.

And if subsequently, we agree again that I shall not cast the fourth throw, I should take $1/6$ of the remainder or $125/1296$ of the total, and I agree with you that that is the value of the fourth throw supposing that one has already made the preceding plays.

But you proposed in the last example in your letter (I quote your very terms) that if I undertake to find the six in eight throws and if I have thrown three times without getting it, and if my opponent

(Facing page 546.)

proposes that I should not play the fourth time, and if he wishes me to be justly treated, it is proper that I have $^{125}/_{1296}$ of the entire sum of our wagers.

This, however, is not true by my theory. For in this case, the three first throws having gained nothing for the player who holds the die, the total sum thus remaining at stake, he who holds the die and who agrees to not play his fourth throw should take $1/6$ as his reward.

And if he has played four throws without finding the desired point and if they agree that he shall not play the fifth time, he will, nevertheless, have $1/6$ of the total for his share. Since the whole sum stays in play it not only follows from the theory, but it is indeed common sense that each throw should be of equal value.

I urge you therefore (to write me) that I may know whether we agree in the theory, as I believe (we do), or whether we differ only in its application.

I am, most heartily, etc.,

Fermat.

Pascal to Fermat
Wednesday, July 29, 1654

Monsieur,—

1. Impatience has seized me as well as it has you, and although I am still abed, I cannot refrain from telling you that I received your letter in regard to the problem of the points[1] yesterday evening from the hands of M. Carcavi, and that I admire it more than I can tell you. I do not have the leisure to write at length, but, in a word, you have found the two divisions of the points and of the dice with perfect justice. I am thoroughly satisfied as I can no longer doubt that I was wrong, seeing the admirable accord in which I find myself with you.

I admire your method for the problem of the points even more than that of the dice. I have seen solutions of the problem of the dice by several persons, as M. le chevalier de Méré, who proposed the question to me, and by M. Roberval also. M. de Méré has

[1] [The editors of these letters note that the word *parti* means the division of the stake between the players in the case when the game is abandoned before its completion. *Parti des dés* means that the man who holds the die agrees to throw a certain number in a given number of trials. For clarity, in this translation, the first of these cases will be called the problem of the points, a term which has had a certain acceptance in the histories of mathematics, while the second may by analogy be called the problem of the dice.]

never been able to find the just value of the problem of the points nor has he been able to find a method of deriving it, so that I found myself the only one who knew this proportion.

2. Your method is very sound and it is the first one that came to my mind in these researches, but because the trouble of these combinations was excessive, I found an abridgment and indeed another method that is much shorter and more neat, which I should like to tell you here in a few words; for I should like to open my heart to you henceforth if I may, so great is the pleasure I have had in our agreement. I plainly see that the truth is the same at Toulouse and at Paris.

This is the way I go about it to know the value of each of the shares when two gamblers play, for example, in three throws, and when each has put 32 pistoles at stake:

Let us suppose that the first of them has *two* (points) and the other *one*. They now play one throw of which the chances are such that if the first wins, he will win the entire wager that is at stake, that is to say 64 pistoles. If the other wins, they will be *two* to *two* and in consequence, if they wish to separate, it follows that each will take back his wager that is to say 32 pistoles.

Consider then, Monsieur, that if the first wins, 64 will belong to him. If he loses, 32 will belong to him. Then if they do not wish to play this point, and separate without doing it, the first should say "I am sure of 32 pistoles, for even a loss gives them to me. As for the 32 others, perhaps I will have them and perhaps you will have them, the risk is equal. Therefore let us divide the 32 pistoles in half, and give me the 32 of which I am certain besides." He will then have 48 pistoles and the other will have 16.

Now let us suppose that the first has *two* points and the other *none*, and that they are beginning to play for a point. The chances are such that if the first wins, he will win all of the wager, 64 pistoles. If the other wins, behold they have come back to the preceding case in which the first has *two* points and the other *one*.

But we have already shown that in this case 48 pistoles will belong to the one who has *two* points. Therefore if they do not wish to play this point, he should say, "If I win, I shall gain all, that is 64. If I lose, 48 will legitimately belong to me. Therefore give me the 48 that are certain to be mine, even if I lose, and let us divide the other 16 in half because there is as much chance that you will gain them as that I will." Thus he will have 48 and 8, which is 56 pistoles.

Let us now suppose that the first has but *one* point and the other *none*. You see, Monsieur, that if they begin a new throw, the chances are such that if the first wins, he will have *two* points to *none*, and dividing by the preceding case, 56 will belong to him. If he loses, they will be point for point, and 32 pistoles will belong to him. He should therefore say, "If you do not wish to play, give me the 32 pistoles of which I am certain, and let us divide the rest of the 56 in half. From 56 take 32, and 24 remains. Then divide 24 in half, you take 12 and I take 12 which with 32 will make 44.

By these means, you see, by simple subtractions that for the first throw, he will have 12 pistoles from the other; for the second, 12 more; and for the last 8.

But not to make this more mysterious, inasmuch as you wish to see everything in the open, and as I have no other object than to see whether I am wrong, the value (I mean the value of the stake of the other player only) of the last play of *two* is double that of the last play of *three* and four times that of the last play of *four* and eight times that of the last play of *five*, etc.

3. But the ratio of the first plays is not so simple to find. This therefore is the method, for I wish to disguise nothing, and here is the problem of which I have considered so many cases, as indeed I was pleased to do: *Being given any number of throws that one wishes, to find the value of the first.*

For example, let the given number of throws be 8. Take the first eight even numbers and the first eight uneven numbers as:

2, 4, 6, 8, 10, 12, 14, 16

and

1, 3, 5, 7, 9, 11, 13, 15.

Multiply the even numbers in this way: the first by the second, their product by the third, their product by the fourth, their product by the fifth, etc.; multiply the odd numbers in the same way: the first by the second, their product by the third, etc.

The last product of the even numbers is the *denominator* and the last product of the odd numbers is the numerator of the fraction that expresses the value of the first throw of *eight*. That is to say that if each one plays the number of pistoles expressed by the product of the even numbers, there will belong to him [who forfeits the throw] the amount of the other's wager expressed by the product of the odd numbers. This may be proved, but with

much difficulty by combinations such as you have imagined, and I have not been able to prove it by this other method which I am about to tell you, but only by that of combinations. Here are the theorems which lead up to this which are properly arithmetic propositions regarding combinations, of which I have found so many beautiful properties:

4. If from any number of letters, as 8 for example,

A, B, C, D, E, F, G, H,

you take all the possible combinations of 4 letters and then all possible combinations of 5 letters, and then of 6, and then of 7, of 8, etc., and thus you would take all possible combinations, I say that if you add together half the combinations of 4 with each of the higher combinations, the sum will be the number equal to the number of the quaternary progression beginning with 2 which is half of the entire number.

For example, and I shall say it in Latin for the French is good for nothing:

If any number whatever of letters, for example 8,

A, B, C, D, E, F, G, H,

be summed in all possible combinations, by fours, fives, sixes, up to eights, I say, if you add half of the combinations by fours, that is 35 (half of 70) to all the combinations by fives, that is 56, and all the combinations by sixes, namely 28, and all the combinations by sevens, namely 8, and all the combinations by eights namely 1, the sum is the fourth number of the quaternary progression whose first term is 2. I say the fourth number for 4 is half of 8.

The numbers of the quaternary progressions whose first term is 2 are

2, 8, 32, 128, 512, etc.,

of which 2 is the first, 8 the second, 32 the third, and 128 the fourth. Of these, the 128 equals:

+ 35 half of the combinations of 4 letters
+ 56 the combinations of 5 letters
+ 28 the combinations of 6 letters
+ 8 the combinations of 7 letters
+ 1 the combinations of 8 letters.

5. That is the first theorem, which is purely arithmetic. The other concerns the theory of the points and is as follows:

It is necessary to say first: if one (player) has *one* point out of 5, for example, and if he thus lacks 4, the game will infallibly be decided in 8 throws, which is double 4.

The value of the first throw of 5 in the wager of the other is the fraction which has for its numerator the half of the combinations of 4 things out of 8 (I take 4 because it is equal to the number of points that he lacks, and 8 because it is double the 4) and for the denominator this same numerator plus all the higher combinations.

Thus if I have one point out of 5, $35/128$ of the wager of my opponent belongs to me. That is to say, if he had wagered 128 pistoles, I would take 35 of them and leave him the rest, 93.

But this fraction $35/128$ is the same as $105/384$, which is made by the multiplication of the even numbers for the denominator and the multiplication of the odd numbers for the numerator.

You will see all of this without a doubt, if you will give yourself a little trouble, and for that reason I have found it unnecessary to discuss it further with you.

6. I shall send you, nevertheless, one of my old Tables; I have not the leisure to copy it, and I shall refer to it.

You will see here as always, that the value of the first throw is equal to that of the second, a thing which may easily be proved by combinations.

You will see likewise that the numbers of the first line are always increasing; those of the second do the same; those of the third the same.

But after that, those of the fourth line diminish; those of the fifth etc. This is odd.

If each wagers 256 on

There belongs to me of the 256 pistoles of my opponent for the	6 throws	5 throws	4 throws	3 throws	2 throws	1 throw
First throw	63	70	80	96	128	256
Second	63	70	80	96	128	
Third	56	60	64	64		
Fourth	42	40	32			
Fifth	24	16				
Sixth	8					

Of the 256 pistoles of my opponent, there belongs to me for the

	If each wagers 256 on					
	6 throws	5 throws	4 throws	3 throws	2 throws	1 throw
First throw	63	70	80	96	128	256
First two throws	126	140	160	192	256	
First three throws	182	200	224	256		
First four throws	224	240	256			
First five throws	248	256				
First six throws	256					

7. I have no time to send you the proof of a difficult point which astonished M. (de Méré) so greatly, for he has ability but he is not a geometer (which is, as you know, a great defect) and he does not even comprehend that a mathematical line is infinitely divisible and he is firmly convinced that it is composed of a finite number of points. I have never been able to get him out of it. If you could do so, it would make him perfect.

He tells me then that he has found an error in the numbers for this reason:

If one undertakes to throw a *six* with a die, the advantage of undertaking to do it in 4 is as 671 is to 625.

If one undertakes to throw double sixes with two dice the disadvantage of the undertaking is 24.

But nonetheless, 24 is to 36 (which is the number of faces of two dice)[1] as 4 is to 6 (which is the number of faces of one die).

This is what was his great scandal which made him say haughtily that the theorems were not consistent and that arithmetic was demented. But you will easily see the reason by the principles which you have.

I shall put all that I have done with this in order when I shall have finished the treatise on geometry[2] on which I have already been working for some time.

8. I have also done something with arithmetic on which subject, I beg you to give me your advice.

---

[1] [Clearly, the number of possible ways in which two dice can fall.]

[2] [Perhaps the manuscript which Leibniz saw, but which is not now extant.]

I proposed the lemma which every one accepts, that the sum of as many numbers as one wishes of the continuous progression from unity as

$$1, 2, 3, 4,$$

being taken by twos is equal to the last term 4 multiplied into the next greater, 5. That is to say that the sum of the integers[1] in A being taken by twos is equal to the product

$$A \times (A + 1).$$

I now come to my theorem:
If one be subtracted from the difference of the cubes of any two consecutive numbers, the result is six times all the numbers contained in the root of the lesser number.

Let the two roots R and S differ by unity. I say that $R^3 - S^3 - 1$ is equal to six times the sum of the numbers contained in S.

Let S be called A, then R is $A + 1$. Therefore the cube of the root R or $A + 1$ is

$$A^3 + 3A^2 + 3A + 1^3.$$

The cube of S, or A, is A^3, and the difference of these is $R^3 - S^3$; therefore, if unity be subtracted, $3A^2 + 3A$ is equal to $R^3 - S^3 - 1$. But by the lemma, double the sum of the numbers contained in A or S is equal to $A \times (A + 1)$; that is, to $A^2 + A$. Therefore, six times the sum of the numbers in A is equal to $3A^3 + 3A$. But $3A^3 + 3A$ is equal to $R^3 - S^3 - 1$. Therefore $R^3 - S^3 - 1$ is equal to six times the sum of the numbers contained in A or S. *Quod erat demonstrandum.* No one has caused me any difficulty in regard to the above, but they have told me that they did not do so for the reason that everyone is accustomed to this method today. As for myself, I mean that without doing me a favor, people should admit this to be an excellent type of proof. I await your comment, however, with all deference. All that I have proved in arithmetic is of this nature.

9. Here are two further difficulties: I have proved a plane theorem making use of the cube of one line compared with the cube of another. I mean that this is purely geometric and in the greatest rigor. By these means I solved the problem: "Any four planes, any four points, and any four spheres being given, to find a sphere which, touching the given spheres, passes through

[1] ["...des nombres contenus dans A."]

the given points, and leaves on the planes segments in which given angles may be inscribed;"[1] and this one: "Any three circles, any three points, and any three lines being given, to find a circle which touches the circles and the points and leaves on the lines an arc in which a given angle may be inscribed."

I solved these problems in a plane, using nothing in the construction but circles and straight lines, but in the proof I made use of solid loci,[2]—of parabolas, or hyperbolas. Nevertheless, inasmuch as the construction is in a plane, I maintain that my solution is plane, and that it should pass as such.

This is a poor recognition of the honor which you have done me in putting up with my discourse which has been plaguing you so long. I never thought I should say two words to you and if I were to tell you what I have uppermost in my heart,—which is that the better I know you the more I honor and admire you,—and if you were to see to what degree that is, you would allot a place in your friendship for him who is, Monsieur, your etc.

Pascal to Fermat
Monday, August 24, 1654

Monsieur,

1. I was not able to tell you my entire thoughts regarding the problem of the points by the last post,[3] and at the same time, I have a certain reluctance at doing it for fear lest this admirable harmony which obtains between us and which is so dear to me should begin to flag, for I am afraid that we may have different opinions on this subject. I wish to lay my whole reasoning before you, and to have you do me the favor to set me straight if I am in error or to indorse me if I am correct. I ask you this in all faith and sincerity for I am not certain even that you will be on my side.

When there are but *two* players, your theory which proceeds by combinations is very just. But when there are three, I believe I have a proof that it is unjust that you should proceed in any other manner than the one I have. But the method which I have disclosed to you and which I have used universally is common to all imaginable conditions of all distributions of points, in the place of that of combinations (which I do not use except in partic-

---

[1] ["...capable d'angles donnés."]

[2] [A common name for conics.]

[3] ["...par l'ordinaire passé." Cf. the English expression, by the "last ordinary."]

ular cases when it is shorter than the general method), a method which is good only in isolated cases and not good for others.

I am sure that I can make it understood, but it requires a few words from me and a little patience from you.

2. This is the method of procedure when there are *two* players: If two players, playing in several throws, find themselves in such a state that the first lacks *two* points and the second *three* of gaining the stake, you say it is necessary to see in how many points the game will be absolutely decided.

It is convenient to suppose that this will be in *four* points, from which you conclude that it is necessary to see how many ways the four points may be distributed between the two players and to see how many combinations there are to make the first win and how many to make the second win, and to divide the stake according to that proportion. I could scarcely understand this reasoning if I had not known it myself before; but you also have written it in your discussion. Then to see how many ways four points may be distributed between two players, it is necessary to imagine that they play with dice with two faces (since there are but two players), as heads and tails, and that they throw four of these dice (because they play in four throws). Now it is necessary to see how many ways these dice may fall. That is easy to calculate. There can be *sixteen*, which is the second power of *four;* that is to say, the square. Now imagine that one of the faces is marked *a*, favorable to the first player. And suppose the other is marked *b*, favorable to the second. Then these four dice can fall according to one of these sixteen arrangements:

a	*a*	*a*	*a*	*a*	*a*	*a*	*a*	*b*	*b*	*b*	*b*	*b*	*b*	*b*	*b*
a	*a*	*a*	*a*	*b*	*b*	*b*	*b*	*a*	*a*	*a*	*a*	*b*	*b*	*b*	*b*
a	*a*	*b*	*b*	*a*	*a*	*b*	*b*	*a*	*a*	*b*	*b*	*a*	*a*	*b*	*b*
a	*b*	*a*	*b*	*a*	*b*	*a*	*b*	*a*	*b*	*a*	*b*	*a*	*b*	*a*	*b*
1	1	1	1	1	1	1	2	1	1	1	2	1	2	2	2

and, because the first player lacks two points, all the arrangements that have two *a*'s make him win. There are therefore 11 of these for him. And because the second lacks three points, all the arrangements that have three *b*'s make him win. There are 5 of these. Therefore it is necessary that they divide the wager as 11 is to 5.

There is your method, when there are *two* players, whereupon you say that if there are more players, it will not be difficult to make the division by this method.

3. On this point, Monsieur, I tell you that this division for the two players founded on combinations is very equitable and good, but that if there are more than two players, it is not always just and I shall tell you the reason for this difference. I communicated your method to [some of] our gentlemen, on which M. de Roberval made me this objection:

That it is wrong to base the method of division on the supposition that they are playing in *four* throws seeing that when one lacks *two* points and the other *three*, there is no necessity that they play *four* throws since it may happen that they play but *two* or *three*, or in truth perhaps *four*.

Since he does not see why one should pretend to make a just division on the assumed condition that one plays *four* throws, in view of the fact that the natural terms of the game are that they do not throw the dice after one of the players has won; and that at least if this is not false, it should be proved. Consequently he suspects that we have committed a paralogism.

I replied to him that I did not found my reasoning so much on this method of combinations, which in truth is not in place on this occasion, as on my universal method from which nothing escapes and which carries its proof with itself. This finds precisely the same division as does the method of combinations. Furthermore, I showed him the truth of the divisions between two players by combinations in this way: Is it not true that if two gamblers finding according to the conditions of the hypothesis that one lacks *two* points and the other *three*, mutually agree that they shall play four complete plays, that is to say, that they shall throw four two-faced dice all at once,—is it not true, I say, that if they are prevented from playing the four throws, the division should be as we have said according to the combinations favorable to each? He agreed with this and this is indeed proved. But he denied that the same thing follows when they are not obliged to play the four throws. I therefore replied as follows:

It is not clear that the same gamblers, not being constrained to play the four throws, but wishing to quit the game before one of them has attained his score, can without loss or gain be obliged to play the whole four plays, and that this agreement in no way changes their condition? For if the first gains the two first points of four, will he who has won refuse to play two throws more, seeing that if he wins he will not win more and if he loses he will not win less? For the two points which the other wins are not sufficient

for him since he lacks three, and there are not enough [points] in four throws for each to make the number which he lacks.

It certainly is convenient to consider that it is absolutely equal and indifferent to each whether they play in the natural way of the game, which is to finish as soon as one has his score, or whether they play the entire four throws. Therefore, since these two conditions are equal and indifferent, the division should be alike for each. But since it is just when they are obliged to play the four throws as I have shown, it is therefore just also in the other case.

That is the way I prove it, and, as you recollect, this proof is based on the equality of the two conditions true and assumed in regard to the two gamblers, the division is the same in each of the methods, and if one gains or loses by one method, he will gain or lose by the other, and the two will always have the same accounting.

4. Let us follow the same argument for *three* players and let us assume that the first lacks *one* point, the second *two*, and the third *two*. To make the division, following the same method of combinations, it is necessary to first discover in how many points the game may be decided as we did when there were two players. This will be in three points for they cannot play three throws without necessarily arriving at a decision.

It is now necessary to see how many ways three throws may be combined among three players and how many are favorable to the first, how many to the second, and how many to the third, and to follow this proportion in distributing the wager as we did in the hypothesis of the two gamblers.

It is easy to see how many combinations there are in all. This is the third power of 3; that is to say, its cube, or 27. For if one throws three dice at a time (for it is necessary to throw three times), these dice having three faces each (since there are three players), one marked *a* favorable to the first, one marked *b* favorable to the second, and one marked *c* favorable to the third,—it is evident that these three dice thrown together can fall in 27 different ways as:

a	*a*	*a*	*a*	*a*	*a*	*a*	*a*	*a*	*b*	*b*	*b*	*b*	*b*	*b*	*b*	*b*	*b*	*c*	*c*	*c*	*c*	*c*	*c*	*c*	*c*	*c*
a	*a*	*a*	*b*	*b*	*b*	*c*	*c*	*c*	*a*	*a*	*a*	*b*	*b*	*b*	*c*	*c*	*c*	*a*	*a*	*a*	*b*	*b*	*b*	*c*	*c*	*c*
a	*b*	*c*	*a*	*b*	*c*	*a*	*b*	*c*	*a*	*b*	*c*	*a*	*b*	*c*	*a*	*b*	*c*	*a*	*b*	*c*	*a*	*b*	*c*	*a*	*b*	*c*
1	1	1	1	1	1	1	1	1	1	1	1	1			1			1	1	1	1			1		
				2						2		2	2	2		2						2				
								3									3			3			3	3	3	3

Since the first lacks but *one* point, then all the ways in which there is one *a* are favorable to him. There are 19 of these. The second lacks *two* points. Thus all the arrangements in which there are two *b*'s are in his favor. There are 7 of them. The third lacks *two* points. Thus all the arrangements in which there are two *c*'s are favorable to him. There are 7 of these.

If we conclude from this that it is necessary to give each according to the proportion 19, 7, 7, we are making a serious mistake and I would hesitate to believe that you would do this. There are several cases favorable to both the first and the second, as *abb* has the *a* which the first needs, and the two *b*'s which the second needs. So too, the *acc* is favorable to the first and third.

It therefore is not desirable to count the arrangements which are common to the two as being worth the whole wager to each, but only as being half a point. For if the arrangement *acc* occurs, the first and third will have the same right to the wager, each making their score. They should therefore divide the wager in half. If the arrangement *aab* occurs, the first alone wins. It is necessary to make this assumption:

There are 13 arrangements which give the entire wager to the first, and 6 which give him half and 8 which are worth nothing to him. Therefore if the entire sum is one pistole, there are 13 arrangements which are each worth one pistole to him, there are 6 that are each worth $\frac{1}{2}$ a pistole, and 8 that are worth nothing.

Then in this case of division, it is necessary to multiply

	13 by one pistole which makes	13
	6 by one half which makes	3
	8 by zero which makes	0
Total 27		Total 16

and to divide the sum of the values 16 by the sum of the arrangements 27, which makes the fraction $^{16}/_{27}$ and it is this amount which belongs to the first gambler in the event of a division; that is to say, 16 pistoles out of 27.

The shares of the second and the third gamblers will be the same:

There are	4 arrangements which are worth 1 pistole; multiplying,	4
There are	3 arrangements which are worth $\frac{1}{2}$ pistole; multiplying,	$1\frac{1}{2}$
And	20 arrangements which are worth nothing	0
Total	27	Total $5\frac{1}{2}$

Therefore $5\frac{1}{2}$ pistoles belong to the second player out of 27, and the same to the third. The sum of the $5\frac{1}{2}$, $5\frac{1}{2}$, and 16 makes 27.

5. It seems to me that this is the way in which it is necessary to make the division by combinations according to your method, unless you have something else on the subject which I do not know. But if I am not mistaken, this division is unjust.

The reason is that we are making a false supposition,—that is, that they are playing three throws without exception, instead of the natural condition of this game which is that they shall not play except up to the time when one of the players has attained the number of points which he lacks, in which case the game ceases.

It is not that it may not happen that they will play three times, but it may happen that they will play once or twice and not need to play again.

But, you will say, why is it possible to make the same assumption in this case as was made in the case of the two players? Here is the reason: In the true condition [of the game] between three players, only one can win, for by the terms of the game it will terminate when one [of the players] has won. But under the assumed conditions, two may attain the number of their points, since the first may gain the one point he lacks and one of the others may gain the two points which he lacks, since they will have played only three throws. When there are only two players, the assumed conditions and the true conditions concur to the advantage of both. It is this that makes the greatest difference between the assumed conditions and the true ones.

If the players, finding themselves in the state given in the hypothesis,—that is to say, if the first lacks *one* point, the second *two*, and the third *two;* and if they now mutually agree and concur in the stipulation that they will play *three* complete throws; and if he who makes the points which he lacks will take the entire sum if he is the only one who attains the points; or if two should attain them that they shall share equally,—*in this case*, the division should be made as I give it here: the first shall have 16, the second $5\frac{1}{2}$, and the third $5\frac{1}{2}$ out of 27 pistoles, and this carries with it its own proof on the assumption of the above condition.

But if they play simply on the condition that they will not necessarily play three throws, but that they will only play until one of them shall have attained his points, and that then the play

shall cease without giving another the opportunity of reaching his score, then 17 pistoles should belong to the first, 5 to the second, and 5 to the third, out of 27. And this is found by my general method which also determines that, under the preceeding condition, the first should have 16, the second 5½, and the third 5½, without making use of combinations,—for this works in all cases and without any obstacle.

6. These, Monsieur, are my reflections on this topic on which I have no advantage over you except that of having meditated on it longer, but this is of little [advantage to me] from your point of view since your first glance is more penetrating than are my prolonged endeavors.

I shall not allow myself to disclose to you my reasons for looking forward to your opinions. I believe you have recognized from this that the theory of combinations is good for the case of two players by accident, as it is also sometimes good in the case of three gamblers, as when one lacks *one* point, another *one*, and the other *two*[1] because, in this case, the number of points in which the game is finished is not enough to allow two to win, but it is not a general method and it is good only in the case where it is necessary to play exactly a certain number of times.

Consequently, as you did not have my method when you sent me the division among several gamblers, but [since you had] only that of combinations, I fear that we hold different views on the subject.

I beg you to inform me how you would proceed in your research on this problem. I shall receive your reply with respect and joy, even if your opinions should be contrary to mine. I am etc.

Fermat to Pascal

Saturday, August 29, 1654

Monsieur,

1. Our interchange of blows still continues, and I am well pleased that our thoughts are in such complete adjustment as it seems since they have taken the same direction and followed the same road. Your recent *Traité du triangle arithmétique*[2] and its applications are an authentic proof and if my computations do

---

[1] [Evidently a misprint, since two throws may be needed.]

[2] [See p. 67.]

me no wrong, your eleventh consequence[1] went by post from Paris to Toulouse while my theorem on figurate numbers,[2] which is virtually the same, was going from Toulouse to Paris.

I have not been on watch for failure while I have been at work on the problem and I am persuaded that the true way to escape failure is by concurring with you. But if I should say more, it would be of the nature of a compliment and we have banished that enemy of sweet and easy conversation.

It is now my turn to give you some of my numerical discoveries, but the end of the parliament augments my duties and I hope that out of your goodness you will allow me due and almost necessary respite.

2. I will reply however to your question of the three players who play in two throws. When the first has one [point] and the others none, your first solution is the true one and the division of the wager should be 17, 5, and 5. The reason for this is self-evident and it always takes the same principle, the combinations making it clear that the first has 17 changes while each of the others has but five.

3. For the rest, there is nothing that I will not write you in the future with all frankness. Meditate however, if you find it convenient, on this theorem: The squared powers of 2 augmented by unity[3] are always prime numbers. [That is,]

The square of 2 augmented by unity makes 5 which is a prime number;

The square of the square makes 16 which, when unity is added, makes 17, a prime number;

The square of 16 makes 256 which, when unity is added, makes 257, a prime number;

The square of 256 makes 65536 which, when unity is added, makes 65537, a prime number;

and so to infinity.

This is a property whose truth I will answer to you. The proof of it is very difficult and I assure you that I have not yet been able to find it fully. I shall not set it for you to find unless I come to the end of it.

---

[1] [From the *Traité du triangle arithmétique,*—"Each cell on the diagonal is double that which preceded it in the parallel or perpendicular rank."]

[2] [*I. e.,* the theorem that $A(A + 1)$ is double the triangular number $1 + 2 + 3 + \ldots A$, See p. 553.]

[3] [*I. e.* $2^{2^n} + 1$. Euler (1732) showed the falsity of the statement.]

This theorem serves in the discovery of numbers which are in a given ratio to their aliquot parts, concerning which I have made many discoveries. We will talk of that another time.

I am Monsieur, yours etc.

Fermat.

At Toulouse, the twenty ninth of August, 1654.

Fermat to Pascal
Friday, September 25, 1654

Monsieur,

1. Do not be apprehensive that our argument is coming to an end. You have strengthened it yourself in thinking to destroy it and it seems to me that in replying to M. de Roberval for yourself you have also replied for me.

In taking the example of the three gamblers of whom the first lacks one point, and each of the others lack two, which is the case in which you oppose, I find here only 17 combinations for the first and 5 for each of the others; for when you say that the combination *acc* is good for the first, recollect that everything that is done after one of the players has won is worth nothing. But this combination having made the first win on the first die, what does it matter that the third gains two afterwards, since even when he gains thirty all this is superfluous? The consequence, as you have well called it "this fiction," of extending the game to a certain number of plays serves only to make the rule easy and (according to my opinion) to make all the chances equal; or better, more intelligibly to reduce all the fractions to the same denomination.

So that you may have no doubt, if instead of *three* parties you extend the assumption to *four*, there will not be 27 combinations only, but 81; and it will be necessary to see how many combinations make the first gain his point later than each of the others gains two, and how many combinations make each of the others win two later than the first wins one. You will find that the combinations that make the first win are 51 and those for each of the other two are 15, which reduces to the same proportion. So that if you take five throws or any other number you please, you will always find three numbers in the proportion of 17, 5, 5. And accordingly I am right in saying that the combination *acc* is [favorable] for the first only and not for the third, and that *cca* is only for the third and not for the first, and consequently my law of combinations is the same for three players as for two, and in general for all numbers.

2. You have already seen from my previous letter that I did not demur at the true solution of the question of the three gamblers for which I sent you the three definite numbers, 17, 5, 5. But because M. de Roberval will perhaps be better satisfied to see a solution without any dissimulation and because it may perhaps yield to abbreviations in many cases, here is an example:

The first may win in a single play, or in two or in three.

If he wins in a single throw, it is necessary that he makes the favorable throw with a three-faced die at the first trial. A single die will yield three chances. The gambler then has $1/3$ of the wager because he plays only one third.

If he plays twice, he can gain in two ways,—either when the second gambler wins the first and he the second, or when the third wins the throw and when he wins the second. But two dice produce 9 chances. The player than has $2/9$ of the wager when they play twice.

But if he plays three times, he can win only in two ways, either the second wins on the first throw and the third wins the second, and he the third; or when the third wins the first throw, the second the second, and he the third; for if the second or the third player wins the two first, he will win the wager and the first player will not. But three dice give 27 chances of which the first player has $2/27$ of the chances when they play three rounds.

The sum of the chances which makes the first gambler win is consequently $1/3$, $2/9$, and $2/27$, which makes $17/27$.

This rule is good and general in all cases of the type where, without recurring to assumed conditions, the true combinations of each number of throws give the solution and make plain what I said at the outset that the extension to a certain number of points is nothing else than the reduction of divers fractions to the same denomination. Here in a few words is the whole of the mystery, which reconciles us without doubt although each of us sought only reason and truth.

3. I hope to send you at Martinmas an abridgment of all that I have discovered of note regarding numbers. You allow me to be concise [since this suffices] to make myself understood to a man [like yourself] who comprehends the whole from half a word. What you will find most important is in regard to the theorem that every number is composed of one, two, or three triangles;[1] of

---

[1] [*I. e.*, triangular numbers.]

one, two, three, or four squares;[1] of one, two, three, four, or five pentagons; of one, two, three, four, five, or six hexagons, and thus to infinity.

To derive this, it is necessary to show that every prime number which is greater by unity than a multiple of 4 is composed of two squares, as 5, 13, 17, 29, 37, etc.

Having given a prime number of this type, as 53, to find by a general rule the two squares which compose it.

Every prime number which is greater by unity than a multiple of 3, is composed of a square and of the triple of another square, as 7, 13, 19, 31, 37, etc.

Every prime number which is greater by 1 or by 3 than a multiple of 8, is composed of a square and of the double of another square, as 11, 17, 19, 41, 43, etc.

There is no triangle of numbers whose area is equal to a square number.

This follows from the invention of many theorems of which Bachet vows himself ignorant and which are lacking in Diophantus.

I am persuaded that as soon as you will have known my way of proof in this type of theorem, it will seem good to you and that it will give you the opportunity for a multitude of new discoveries, for it follows as you know that *multi pertranseant ut augeatur scientia.*

When I have time, we will talk further of magic numbers and I will summarize my former work on this subject.

I am, Monsieur, most heartily your etc.

Fermat.

The twenty-fifth of September.

I am writing this from the country, and this may perhaps delay my replies during the holidays.

Pascal to Fermat

Tuesday, October 27, 1654

Monsieur,

Your last letter satisfied me perfectly. I admire your method for the problem of the points, all the more because I understand it well. It is entirely yours, it has nothing in common with mine, and it reaches the same end easily. Now our harmony has begun again.

But, Monsieur, I agree with you in this, find someone elsewhere to follow you in your discoveries concerning numbers, the state-

[1] [See page 91.]

ments of which you were so good as to send me. For my own part, I confess that this passes me at a great distance; I am competent only to admire it and I beg you most humbly to use your earliest leisure to bring it to a conclusion. All of our gentlemen saw it on Saturday last and appreciate it most heartily. One cannot often hope for things that are so fine and so desirable. Think about it if you will, and rest assured that I am etc.

Pascal.

Paris, October 27, 1654.

DE MOIVRE

On the Law of Normal Probability

(Edited by Professor Helen M. Walker, Teachers College, Columbia University, New York City.)

Abraham de Moivre (1667–1754) left France at the revocation of the Edict of Nantes and spent the rest of his life in London, where he solved problems for wealthy patrons and did private tutoring in mathematics. He is best known for his work on trigonometry, probability, and annuities. On November 12, 1733 he presented privately to some friends a brief paper of seven pages entitled "Approximatio ad Summam Terminorum Binomii $\overline{a + b}|^n$ in Seriem expansi." Only two copies of this are known to be extant. His own translation with some additions, was included in the second edition (1738) of *The Doctrine of Chances*, pages 235–243.

This paper gave the first statement of the formula for the "normal curve," the first method of finding the probability of the occurrence of an error of a given size when that error is expressed in terms of the variability of the distribution as a unit, and the first recognition of that value later termed the *probable error*. It shows, also, that before Stirling, De Moivre had been approaching a solution of the value of factorial n.

A Method of approximating the Sum of the Terms of the Binomial $\overline{a+b}|^n$ expanded into a Series, from whence are deduced some practical Rules to estimate the Degree of Assent which is to be given to Experiments.

Altho' the Solution of Problems of Chance often require that several Terms of the Binomial $\overline{a+b}|^n$ be added together, nevertheless in very high Powers the thing appears so laborious, and of so great a difficulty, that few people have undertaken that Task; for besides *James* and *Nicolas Bernoulli*, two great Mathematicians, I know of no body that has attempted it; in which, tho' they have shewn very great skill, and have the praise which is due to their Industry, yet some things were farther required; for what they have done is not so much an Approximation as the determining very wide limits, within which they demonstrated that the Sum of the Terms was contained. Now the Method which they have followed has been briefly described in my *Miscellanea Analytica*, which the Reader may consult if he pleases,

unless they rather chuse, which perhaps would be the best, to consult what they themselves have writ upon that Subject: for my part, what made me apply myself to that Inquiry was not out of opinion that I should excel others, in which however I might have been forgiven; but what I did was in compliance to the desire of a very worthy Gentleman, and good Mathematician, who encouraged me to it: I now add some new thoughts to the former; but in order to make their connexion the clearer, it is necessary for me to resume some few things that have been delivered by me a pretty while ago.

I. It is now a dozen years or more since I had found what follows; If the Binomial $1+1$ be raised to a very high Power denoted by n, the ratio which the middle Term has to the Sum of all the Terms, that is, to 2^n, may be expressed by the Fraction $\frac{2A\times\overline{n-1}|^n}{n^n\times\sqrt{n-1}}$, wherein A represents the number of which the Hyperbolic Logarithm is $\frac{1}{12}-\frac{1}{360}+\frac{1}{1260}-\frac{1}{1680}$, &c. but because the Quantity $\frac{\overline{n-1}|^n}{n^n}$ or $\overline{1-\frac{1}{n}}|^n$ is very nearly given when n is a high Power, which is not difficult to prove, it follows that, in an infinite Power, that Quantity will be absolutely given, and represent the number of which the Hyperbolic Logarithm is -1; from whence it follows, that if B denotes the Number of which the Hyperbolic Logarithm is $-1+\frac{1}{12}-\frac{1}{360}+\frac{1}{1260}-\frac{1}{1680}$, &c. the Expression above-written will become $\frac{2B}{\sqrt{n-1}}$ or barely $\frac{2B}{\sqrt{n}}$, and that therefore if we change the Signs of that Series, and now suppose that B represents the Number of which the Hyperbolic Logarithm is $1-\frac{1}{12}+\frac{1}{360}-\frac{1}{1260}+\frac{1}{1680}$, &c. that Expression will be changed into $\frac{2}{B\sqrt{n}}$.

When I first began that inquiry, I contented myself to determine at large the Value of B, which was done by the addition of some Terms of the above-written Series; but as I perceiv'd that it converged but slowly, and seeing at the same time that what I had done answered my purpose tolerably well, I desisted from proceeding farther, till my worthy and learned Friend Mr. *James Stirling*, who had applied himself after me to that inquiry, found

that the Quantity B did denote the Square-root of the Circumference of a Circle whose Radius is Unity, so that if that Circumference be called c, the Ratio of the middle Term to the Sum of all the Terms will be expressed by $\frac{2}{\sqrt{nc}}$.[1]

But altho' it be not necessary to know what relation the number B may have to the Circumference of the Circle, provided its value be attained, either by pursuing the Logarithmic Series before mentioned, or any other way; yet I own with pleasure that this discovery, besides that it has saved trouble, has spread a singular Elegancy on the Solution.

II. I also found that the Logarithm of the Ratio which the middle Term of a high Power has to any Term distant from it by an Interval denoted by l, would be denoted by a very near approximation, (supposing $m = \frac{1}{2}n$) by the Quantities $\overline{m+l-\frac{1}{2}}\times\log.\overline{m+l-1} + \overline{m-l+\frac{1}{2}}\times\log.\overline{m-l+1} - 2m\times\log. m + \log.\frac{m+l}{m}$.

Corollary 1.

This being admitted, I conclude, that if m or $\frac{1}{2}n$ be a Quantity infinitely great, then the Logarithm of the Ratio, which a Term distant from the middle by the Interval l, has to the middle Term, is $-\frac{2ll}{n}$.[2]

Corollary 2.

The Number, which answers to the Hyperbolic Logarithm $-\frac{2ll}{n}$, being

$$1-\frac{2ll}{n}+\frac{4l^4}{2nn}-\frac{8l^6}{6n^3}+\frac{16l^8}{24n^4}-\frac{32l^{10}}{120n^5}+\frac{64l^{12}}{720n^6}, \text{\&c.}$$

---

[1] [Under the circumstances of De Moivre's problem, nc is equivalent to $8\sigma^2\pi$, where σ is the standard deviation of the curve. This statement therefore shows that De Moivre knew the maximum ordinate of the curve to be

$$y_0 = \frac{1}{\sigma\sqrt{2\pi}}.]$$

[2] [Since $n = 4\sigma^2$ under the assumptions made here, this is equivalent to stating the formula for the curve as

$$y = y_0 e^{-\frac{x^2}{2\sigma^2}}.]$$

it follows, that the Sum of the Terms intercepted between the Middle, and that whose distance from it is denoted by l, will be

$$\frac{2}{\sqrt{nc}} \text{ into } l - \frac{2l^3}{1\times 3n} + \frac{4l^5}{2\times 5nn} - \frac{8l^7}{6\times 7n^3} + \frac{16l^9}{24\times 9n^4} - \frac{32l^{11}}{120\times 11n^5}, \text{ \&c.}$$

Let now l be supposed $= s\sqrt{n}$, then the said Sum will be expressed by the Series

$$\frac{2}{\sqrt{c}} \text{ into } ʃ - \frac{2ʃ^3}{3} + \frac{4ʃ^5}{2\times 5} - \frac{8ʃ^7}{6\times 7} + \frac{16ʃ^9}{24\times 9} - \frac{32ʃ^{11}}{120\times 11}, \text{ \&c.}$$ [1]

Moreover, if $ʃ$ be interpreted by $\frac{1}{2}$, then the Series will become

$$\frac{2}{\sqrt{c}} \text{ into } \frac{1}{2} - \frac{1}{3\times 4} + \frac{1}{2\times 5\times 8} - \frac{1}{6\times 7\times 16} + \frac{1}{24\times 9\times 32} - \frac{1}{120\times 11\times 64},$$

&c. which converges so fast, that by help of no more than seven or eight Terms, the Sum required may be carried to six or seven places of Decimals: Now that Sum will be found to be 0.427812, independently from the common Multiplicator $\frac{2}{\sqrt{c}}$, and therefore to the Tabular Logarithm of 0.427812, which is $\overline{9}.6312529$, adding the Logarithm of $\frac{2}{\sqrt{c}}$, *viz.* $\overline{9}.9019400$, the Sum will be $\overline{19}.5331929$, to which answers the number 0.341344.

LEMMA.

If an Event be so dependent on Chance, as that the Probabilities of its happening or failing be equal, and that a certain given number n of Experiments be taken to observe how often it happens and fails, and also that l be another given number, less than $\frac{1}{2}n$, then the Probability of its neither happening more frequently than $\frac{1}{2}n+l$ times, nor more rarely than $\frac{1}{2}n-l$ times, may be found as follows.

Let L and L be two Terms equally distant on both sides of the middle Term of the Binomial $\overline{1+1}|^n$ expanded, by an Interval equal to l; let also $ʃ$ be the Sum of the Terms included between L and L together with the Extreams, then the Probability required will be rightly expressed by the Fraction $\frac{ʃ}{2^n}$, which being founded on the common Principles of the Doctrine of Chances, requires no Demonstration in this place.

[1] [The long *s* which De Moivre employed in this formula is not to be mistaken for the integral sign.]

Corollary 3.

And therefore, if it was possible to take an infinite number of Experiments, the Probability that an Event which has an equal number of Chances to happen or fail, shall neither appear more frequently than $\frac{1}{2}n + \frac{1}{2}\sqrt{n}$ times, nor more rarely than $\frac{1}{2}n - \frac{1}{2}n\sqrt{n}$ times, will be express'd by the double Sum of the number exhibited in the second Corollary, that is, by 0.682688, and consequently the Probability of the contrary, which is that of happening more frequently or more rarely than in the proportion above assigned will be 0.317312, those two Probabilities together compleating Unity, which is the measure of Certainty: Now the Ratio of those Probabilities is in small Terms 28 to 13 very near.

Corollary 4.

But altho' the taking an infinite number of Experiments be not practicable, yet the preceding Conclusions may very well be applied to finite numbers, provided they be great, for Instance, if 3600 Experiments be taken, make $n = 3600$, hence $\frac{1}{2}n$ will be $= 1800$, and $\frac{1}{2}\sqrt{n} = 30$, then the Probability of the Event's neither appearing oftner than 1830 times, nor more rarely than 1770, will be 0.682688.

Corollary 5.

And therefore we may lay this down for a fundamental Maxim, that in high Powers, the Ratio, which the Sum of the Terms included between two Extreams distant on both sides from the middle Term by an Interval equal to $\frac{1}{2}\sqrt{n}$, bears to the Sum of all the Terms, will be rightly express'd by the Decimal 0.682688, that is $\frac{28}{41}$ nearly.

Still, it is not to be imagin'd that there is any necessity that the number n should be immensely great; for supposing it not to reach beyond the 900^{th} Power, nay not even beyond the 100^{th}, the Rule here given will be tolerably accurate, which I have had confirmed by Trials.

But it is worth while to observe, that such a small part as is $\frac{1}{2}\sqrt{n}$ in respect to n, and so much the less in respect to n as n increases, does very soon give the Probability $\frac{28}{41}$ or the Odds of 28 to 13; from whence we may naturally be led to enquire, what are the

Bounds within which the proportion of Equality is contained; I answer, that these Bounds will be set at such a distance from the middle Term, as will be expressed by $\frac{1}{4}\sqrt{2n}$ very near; so in the case above mentioned, wherein n was supposed $=3600$, $\frac{1}{4}\sqrt{2n}$ will be about 21.2 nearly, which in respect to 3600, is not above $\frac{1}{169}$th part: so that it is an equal Chance nearly, or rather something more, that in 3600 Experiments, in each of which an Event may as well happen as fail, the Excess of the happenings or failings above 1800 times will be no more than about 21.

Corollary 6.

If l be interpreted by $\sqrt{n}$, the Series will not converge so fast as it did in the former case when l was interpreted by $\frac{1}{2}\sqrt{n}$, for here no less than 12 or 13 Terms of the Series will afford a tolerable approximation, and it would still require more Terms, according as l bears a greater proportion to $\sqrt{n}$, for which reason I make use in this case of the Artifice of Mechanic Quadratures, first invented by Sir *Isaac Newton*, and since prosecuted by Mr. *Cotes*, Mr. *James Stirling*, myself, and perhaps others; it consists in determining the Area of a Curve nearly, from knowing a certain number of its Ordinates A, B, C, D, E, F, &c. placed at equal Intervals, the more Ordinates there are, the more exact will the Quadrature be; but here I confine myself to four, as being sufficient for my purpose: let us therefore suppose that the four Ordinates are A, B, C, D, and that the Distance between the first and last is denoted by l, then the Area contained between the first and the last will be $\frac{1\times\overline{A+D}+3\times\overline{B+C}}{8}\times l$; now let us take the Distances $0\sqrt{n}$, $\frac{1}{6}\sqrt{n}$, $\frac{2}{6}\sqrt{n}$, $\frac{3}{6}\sqrt{n}$, $\frac{4}{6}\sqrt{n}$, $\frac{5}{6}\sqrt{n}$, $\frac{6}{6}\sqrt{n}$, of which every one exceeds the preceding by $\frac{1}{6}\sqrt{n}$, and of which the last is $\sqrt{n}$; of these let us take the four last, *viz.* $\frac{3}{6}\sqrt{n}$, $\frac{4}{6}\sqrt{n}$, $\frac{5}{6}\sqrt{n}$, $\frac{6}{6}\sqrt{n}$, then taking their Squares, doubling each of them, dividing them all by n, and prefixing to them all the Sign $-$, we shall have $-\frac{1}{2}$, $-\frac{8}{9}$, $-\frac{25}{18}$, $-\frac{2}{1}$, which must be look'd upon as Hyperbolic Logarithms, of which consequently the corresponding numbers, *viz.* 0.60653, 0.41111, 0.24935, 0.13534 will stand for the four Ordinates A, B, C, D.

Now having interpreted l by $\frac{1}{2}\sqrt{n}$, the Area will be found to be $=0.170203\times\sqrt{n}$, the double of which being multiplied by $\frac{2}{\sqrt{nc}}$, the product will be 0.27160; let therefore this be added to the Area found before, that is, to 0.682688, and the Sum 0.95428 will shew, what after a number of Trials denoted by n, the Probability will be of the Event's neither happening oftner than $\frac{1}{2}n+\sqrt{n}$ times, nor more rarely than $\frac{1}{2}n-\sqrt{n}$, and therefore the Probability of the contrary will be 0.04572, which shews that the Odds of the Event's neither happening oftner nor more rarely than within the Limits assigned are 21 to 1 nearly.

And by the same way of reasoning, it will be found that the Probability of the Event's neither appearing oftner $\frac{1}{2}n+\frac{3}{2}\sqrt{n}$, nor more rarely than $\frac{1}{2}n-\frac{3}{2}\sqrt{n}$ will be 0.99874, which will make it that the Odds in this case will be 369 to 1 nearly.

To apply this to particular Examples, it will be necessary to estimate the frequency of an Event's happening or failing by the Square-root of the number which denotes how many Experiments have been, or are designed to be taken; and this Square-root, according as it has been already hinted at in the fourth Corollary, will be as it were the Modulus by which we are to regulate our Estimation; and therefore suppose the number of Experiments to be taken is 3600, and that it were required to assign the Probability of the Event's neither happening oftner than 2850 times, nor more rarely than 1750, which two numbers may be varied at pleasure, provided they be equally distant from the middle Sum 1800, then make the half difference between the two numbers 1850 and 1750, that is, in this case, $50=ʃ\sqrt{n}$; now having supposed $3600=n$, then $\sqrt{n}$ will be $=60$, which will make it that 50 will be $=60ʃ$, and consequently $ʃ=\frac{50}{60}=\frac{5}{6}$; and therefore if we take the proportion, which in an infinite power, the double Sum of the Terms corresponding to the Interval $\frac{5}{6}\sqrt{n}$, bears to the Sum of all the Terms, we shall have the Probability required exceeding near.

Lemma 2.

In any Power $\overline{a+b}^{n}$ expanded, the greatest Term is that in which the Indices of the Powers of a and b, have the same propor-

tion to one another as the Quantities themselves a and b; thus taking the 10^{th} Power of $a+b$, which is $a^{10}+10a^9b+45a^8b^2+120a^7b^3+210a^6b^4+252a^5b^5+210a^4b^6+120a^3b^7+45a^2b^8+10ab^9+b^{10}$; and supposing that the proportion of a to b is as 3 to 2, then the Term $210a^6b^4$ will be the greatest, by reason that the Indices of the Powers of a and b, which are in that Term, are in the proportion of 3 to 2; but supposing the proportion of a to b had been as 4 to 1, then the Term $45a^8b^2$ had been the greatest.

Lemma 3.

If an Event so depends on Chance, as that the Probabilities of its happening or failing be in any assigned proportion, such as may be supposed of a to b, and a certain number of Experiments be designed to be taken, in order to observe how often the Event will happen or fail; then the Probability that it shall neither happen more frequently than so many times as are denoted by $\frac{an}{a+b}+l$, nor more rarely than so many times as are denoted by $\frac{an}{a+b}-l$, will be found as follows:

Let L and R be equally distant by the Interval l from the greatest Term; let also S be the Sum of the Terms included between L and R, together with those Extreams, then the Probability required will be rightly expressed by $\frac{S}{\overline{a+b}^{\,n}}$.

Corollary 8.[1]

The Ratio which, in an infinite Power denoted by n, the greatest Term bears to the Sum of all the rest, will be rightly expressed by the Fraction $\frac{a+b}{\sqrt{abnc}}$, wherein c denotes, as before, the Circumference of a Circle for a Radius equal to Unity.

Corollary 9.

If, in an infinite Power, any Term be distant from the Greatest by the Interval l, then the Hyperbolic Logarithm of the Ratio which that Term bears to that Greatest will be expressed by the Fraction $-\frac{\overline{a+b}^{\,2}}{2abn}\times ll$; provided the Ratio of l to n be not a finite

---

[1] [Numbered as in the original. There is no corollary 7 in the text.]

Ratio, but such a one as may be conceived between any given number p and $\sqrt{n}$, so that l be expressible by $p\sqrt{n}$, in which case the two Terms L and R will be equal.

COROLLARY 10.

If the Probabilities of happening and failing be in any given Ratio of unequality, the Problems relating to the Sum of the Terms of the Binomial $\overline{a+b}|^n$ will be solved with the same facility as those in which the Probabilities of happening and failing are in a Ratio of Equality.

From what has been said, it follows, that Chance very little disturbs the Events which in their natural Institution were designed to happen or fail, according to some determinate Law; for if in order to help our conception, we imagine a round piece of Metal, with two polished opposite faces, differing in nothing but their colour, whereof one may be supposed to be white, and the other black; it is plain that we may say, that this piece may with equal facility exhibit a white or black face, and we may even suppose that it was framed with that particular view of shewing sometimes one face, sometimes the other, and that consequently if it be tossed up Chance shall decide the appearance; but we have seen in our LXXXVII[th] Problem, that altho' Chance may produce an inequality of appearance, and still a greater inequality according to the length of time in which it may exert itself, still the appearances, either one way or the other, will perpetually tend to a proportion of Equality; but besides we have seen in the present Problem, that in a great number of Experiments, such as 3600, it would be the Odds of above 2 to 1, that one of the Faces, suppose the white, shall not appear more frequently than 1830 times, nor more rarely than 1770, or in other Terms, that it shall not be above or under the perfect Equality by more than $\frac{1}{120}$ part of the whole number of appearances; and by the same Rule, that if the number of Trials had been 14400 instead of 3600, then still it would be above the Odds of 2 to 1, that the appearances either one way or other would not deviate from perfect Equality by more than $\frac{1}{260}$ part of the whole, but in 1000000 Trials it would be the Odds of above 2 to 1, that the deviation from perfect Equality would not be more than by $\frac{1}{2000}$ part of the whole. But the

Odds would increase at a prodigious rate, if instead of taking such narrow limits on both sides the Term of Equality, as are represented by $\frac{1}{2}\sqrt{n}$, we double those Limits or triple them; for in the first case the Odds would become 21 to 1, and in the second 369 to 1, and still be vastly greater if we were to quadruple them, and at last be infinitely great; and yet whether we double, triple or quadruple them, &c. the Extension of those Limits will bear but an inconsiderable proportion to the whole, and none at all, if the whole be infinite, of which the reason will easily be perceived by Mathematicians, who know, that the Square-root of any Power bears so much a less proportion to that Power, as the Index of it is great.

And what we have said is also applicable to a Ratio of Inequality, as appears from our 9th Corollary. And thus in all cases it will be found, that altho' Chance produces irregularities, still the Odds will be infinitely great, that in process of Time, those Irregularities will bear no proportion to the recurrency of that Order which naturally results from original Design.

LEGENDRE

On Least Squares

(Translated from the French by Professor Henry A. Ruger and Professor Helen M. Walker, Teachers College, Columbia University, New York City.)

The great advances in mathematical astronomy made during the early year) of the nineteenth century were due in no small part to the development of the method of least squares. The same method is the foundation for the calculus of errors of observation now occupying a place of great importance in the scientific study of social, economic, biological, and psychological problems. Gauss says in his work on the *Theory of the Motions of the Heavenly Bodies* (1809) that he had made use of this principle since 1795 but that it was first published by Legendre. The first statement of the method appeared as an appendix entitled "Sur la Méthode des moindres quarrés" in Legendre's *Nouvelles méthodes pour la détermination des orbites des comètes*, Paris, 1805. The portion of the work translated here is found on pages 72–75.

Adrien-Marie Legendre (1752–1833) was for five years a professor of mathematics in the École Militaire at Paris, and his early studies on the paths of projectiles provided a background for later work on the paths of heavenly bodies. He wrote on astronomy, the theory of numbers, elliptic functions, the calculus, higher geometry, mechanics, and physics. His work on geometry, in which he rearranged the propositions of Euclid, is one of the most successful textbooks ever written.

On the Method of Least Squares

In the majority of investigations in which the problem is to get from measures given by observation the most exact results which they can furnish, there almost always arises a system of equations of the form

$$E = a + bx + cy + fz + \&c.$$

in which a, b, c, f, &c. are the known coefficients which vary from one equation to another, and x, y, z, &c. are the unknowns which must be determined in accordance with the condition that the value of E shall for each equation reduce to a quantity which is either zero or very small.

If there are the same number of equations as unknowns x, y, z, &c., there is no difficulty in determining the unknowns, and the error E can be made absolutely zero. But more often the number

of equations is greater than that of the unknowns, and it is impossible to do away with all the errors.

In a situation of this sort, which is the usual thing in physical and astronomical problems, where there is an attempt to determine certain important components, a degree of arbitrariness necessarily enters in the distribution of the errors, and it is not to be expected that all the hypotheses shall lead to exactly the same results; but it is particularly important to proceed in such a way that extreme errors, whether positive or negative, shall be confined within as narrow limits as possible.

Of all the principles which can be proposed for that purpose, I think there is none more general, more exact, and more easy of application, than that of which we have made use in the preceding researches, and which consists of rendering the sum of the squares of the errors a *minimum*. By this means there is established among the errors a sort of equilibrium which, preventing the extremes from exerting an undue influence, is very well fitted to reveal that state of the system which most nearly approaches the truth.

The sum of the squares of the errors $E^2 + E'^2 + E''^2 +$ &c. being

$$\begin{aligned} &(a + bx + cy + fz + \&c.)^2 \\ + &(a' + b'x + c'y + f'z + \&c.)^2 \\ + &(a'' + b''x + c''y + f''z + \&c.)^2 \\ + &\&c., \end{aligned}$$

if its *minimum* is desired, when x alone varies, the resulting equation will be

$$o = \int ab + x\int b^2 + y\int bc + z\int bf + \&c.,$$

in which by $\int ab$ we understand the sum of similar products, *i.e.*, $ab + a'b' + a''b'' +$ &c.; by $\int b^2$ the sum of the squares of the coefficients of x, namely $b^2 + b'^2 + b''^2 +$ &c., and similarly for the other terms.

Similarly the *minimum* with respect to y will be

$$o = \int ac + x\int bc + y\int c^2 + z\int fc + \&c.,$$

and the *minimum* with respect to z,

$$o = \int af + x\int bf + y\int cf + z\int f^2 + \&c.,$$

in which it is apparent that the same coefficients $\int bc$, $\int bf$, &c. are common to two equations, a fact which facilitates the calculation.

In general, *to form the equation of the* minimum *with respect to one of the unknowns, it is necessary to multiply all the terms of each given equation by the coefficient of the unknown in that equation, taken with regard to its sign, and to find the sum of these products.*

The number of equations of *minimum* derived in this manner will be equal to the number of the unknowns, and these equations are then to be solved by the established methods. But it will be well to reduce the amount of computation both in multiplication and in solution, by retaining in each operation only so many significant figures, integers or decimals, as are demanded by the degree of approximation for which the inquiry calls.

Even if by a rare chance it were possible to satisfy all the equations at once by making all the errors zero, we could obtain the same result from the equations of *minimum;* for if after having found the values of x, y, z, &c. which make E, E', &c. equal to zero, we let x, y, z, &c. vary by δx, δy, δz, &c., it is evident that E^2, which was zero, will become by that variation $(a\delta x + b\delta y + c\delta z + \text{\&c.})^2$. The same will be true of E'^2, E''^2, &c. Thus we see that the sum of the squares of the errors will by variation become a quantity of the second order with respect to δx, δy, &c., which is in accord with the nature of a minimum.

If after having determined all the unknowns x, y, z, &c., we substitute their values in the given equations, we will find the value of the different errors E, E', E'', &c., to which the system gives rise, and which cannot be reduced without increasing the sum of their squares. If among these errors are some which appear too large to be admissible, then those equations which produced these errors will be rejected, as coming from too faulty experiments, and the unknowns will be determined by means of the other equations, which will then give much smaller errors. It is further to be noted that one will not then be obliged to begin the calculations anew, for since the equations of minimum are formed by the addition of the products made in each of the given equations, it will suffice to remove from the addition those products furnished by the equations which would have led to errors that were too large.

The rule by which one finds the mean among the results of different observations is only a very simple consequence of our general method, which we will call *the method of least squares.*

Indeed, if experiments have given different values a', a'', a''', &c. for a certain quantity x, the sum of the squares of the errors

will be $(a' - x)^2 + (a'' - x)^2 + (a''' - x)^2 + \&c.$, and on making that sum a minimum, we have

$$o = (a' - x) + (a'' - x) + (a''' - x) + \&c.,$$

from which it follows that

$$x = \frac{a' + a'' + a''' + \&c.}{n},$$

n being the number of the observations.

In the same way, if to determine the position of a point in space, a first experiment has given the coordinates a', b', c'; a second, the coordinates a'', b'', c''; and so on, and if the true coordinates of the point are denoted by x, y, z; then the error in the first experiment will be the distance from the point (a', b', c') to the point (x, y, z). The square of this distance is

$$(a' - x)^2 + (b' - y)^2 + (c' - z)^2.$$

If we make the sum of the squares of all such distances a minimum, we get three equations which give

$$x = \frac{\int a}{n}, \quad y = \frac{\int b}{n}, \quad z = \frac{\int c}{n},$$

n being the number of points given by the experiments. These formulas are precisely the ones by which one might find the common center of gravity of several equal masses situated at the given points, whence it is evident that the center of gravity of any body possesses this general property.

If we divide the mass of a body into particles which are equal and sufficiently small to be treated as points, the sum of the squares of the distances from the particles to the center of gravity will be a minimum.

We see then that the method of least squares reveals to us, in a fashion, the center about which all the results furnished by experiments tend to distribute themselves, in such a manner as to make their deviations from it as small as possible. The application which we are now about to make of this method to the measurement of the meridian will display most clearly its simplicity and fertility.[1]

[1] [An application of the method to an astronomical problem follows.]

CHEBYSHEV (Tchebycheff)

Theorem Concerning Mean Values

(Translated from the French by Professor Helen M. Walker, Teachers College, Columbia University, New York City.)

The inequality which Chebyshev (Tchebycheff) derived in his paper on mean values is an important contribution to the theory of dispersion. In this paper by simple algebra, without approximation or the aid of the calculus, he reached a result from which both "Jacques Bernoulli's Theorem" and Poisson's "Law of Large Numbers" can be derived as special cases. The selection printed here was translated from the Russian into French by M. N. de Khanikof and appeared in Liouville's *Journal de mathématiques pures et appliquées ou recueil mensuel de mémoires sur les diverses parties des mathématiques*, 2nd series, XII (1867), 177–184. The same material is to be found in his *Œuvres*, I (1899).

Pafnutii Lvovitch Chebyshev (Tchebysheff)[1] (1821–1894) was, after Lobachevsky, Russia's most celebrated mathematician. Even as a small boy he was greatly interested in mechanical inventions, and it is said that in his first lesson in geometry he saw the bearing of the subject upon mechanics and therefore resolved to master it. At the age of twenty he received his diploma from the University of Moscow, having already received a medal for a work on the numerical solution of algebraic equations of higher orders.

Chebyshev's father was a Russian nobleman, but after the famine of 1840 the estate was so reduced that for the rest of his life he was forced to practice extreme economy, spending money freely for nothing except the mechanical models of his various inventions. He never married, but devoted himself solely to science.

Chebyshev collaborated with Bouniakovsky in bringing out the two large volumes of the collected works of Euler in 1849 and this seems to have turned his thoughts to the theory of numbers, and particularly to the very difficult problem of the distribution of primes. In 1850 he established the existence of limits within which must be comprised the sum of the logarithms of primes inferior to a given number. In 1860 he was made a correspondent of the Institut de France, and in 1874 an *associé étranger*. He was also a foreign member of the Royal Society of London.

From 1847 to 1882 he was professor of mathematics at the University of St. Petersburg, and at different periods during this time he taught analytic

[1] The name is spelled in many ways such as Chebychef, Chebichev, Tchebychev, Tchébycheff, Tschebycheff. For further biographical details the reader is referred to a brochure by A. Vassilief entitled *P. L. Tchébychef et son Œuvre Scientifique* (Turin, 1898) reprinted from the *Bollettino di bibliografia e storia delle scienze matematiche pubblicato per cura di Gino Loria*, 1898, or to the sketch by C. A. Possé in the *Dictionnaire des écrivains et savants russes rédigé par M. Venguerof*, reprinted in Volume II of Markof's edition of Chebyshev's *Œuvres*.

geometry, higher algebra, theory of numbers, integral calculus, theory of probabilities, the calculus of finite differences, the theory of elliptic functions, and the theory of definite integrals, and his biographers are agreed that the quality of his teaching was no less remarkable than that of his research. Chebyshev made important contributions to the theory of numbers, theory of least squares, interpolation theory, calculus of variations, infinite series, and the theory of probability, and published nearly a hundred memoirs on these and other mathematical topics, being best known for his work on primes. The very day before his death he received his friends as usual and discoursed upon the subject of a simple rule he had discovered for the rectification of a curve.

On the Mean Values

If we agree to speak of the *mathematical expectation* of any magnitude as the sum of all the values which it may assume multiplied by their respective probabilities, it will be easy for us to establish a very simple theorem concerning the limits between which shall be contained a sum of any values whatever.

THEOREM. *If we designate by* $a, b, c \ldots,$ *the mathematical expectations of the quantities* $x, y, z \ldots,$ *and by* $a_1, b_1, c_1 \ldots,$ *the mathematical expectations of their squares* $x^2, y^2, z^2 \ldots,$ *the probability that the sum* $x + y + z \ldots$ *is included within the limits*

$$a + b + c + \ldots + \alpha\sqrt{a_1 + b_1 + c_1 + \ldots - a^2 - b^2 - c^2 - \ldots},$$

and

$$a + b + c + \ldots - \alpha\sqrt{a_1 + b_1 + c_1 + \ldots - a^2 - b^2 - c^2 - \ldots},$$

will always be larger than $1 - \frac{1}{\alpha^2}$, *no matter what the size of* α.

Proof. Let

$$\begin{array}{l} x_1,\ x_2,\ x_3, \ldots,\ x_l, \\ y_1,\ y_2,\ y_3, \ldots,\ y_m, \\ z_1,\ z_2,\ z_3, \ldots\ z_n, \\ \ldots\ldots\ldots\ldots\ldots \end{array}$$

be all conceivable values of the quantities $x, y, z, \ldots,$ and let

$$\begin{array}{l} p_1,\ p_2,\ p_3, \ldots,\ p_l, \\ q_1,\ q_2,\ q_3, \ldots,\ q_m, \\ r_1,\ r_2,\ r_3, \ldots,\ r_n, \\ \ldots\ldots\ldots\ldots\ldots \end{array}$$

be the respective probabilities of these values, or, better, the probabilities of the hypotheses

$$\begin{array}{l} x = x_1, x_2, x_3, \ldots, x_l, \\ y = y_1, y_2, y_3, \ldots, y_m, \\ z = z_1, z_2, z_3, \ldots, z_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots \end{array}$$

In accordance with this notation, the mathematical expectations of the magnitudes $x, y, z, \ldots$, and of $x^2, y^2, z^2, \ldots$ will be expressed as follows:

$$(1) \qquad \begin{cases} a = p_1x_1 + p_2x_2 + p_3x_3 + \ldots + p_lx_l, \\ b = q_1y_1 + q_2y_2 + q_3y_3 + \ldots + q_my_m, \\ c = r_1z_1 + r_2z_2 + r_3z_3 + \ldots + r_nz_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{cases}$$

$$(2) \qquad \begin{cases} a_1 = p_1x_1^2 + p_2x_2^2 + p_3x_3^2 + \ldots + p_lx_l^2, \\ b_1 = q_1y_1^2 + q_2y_2^2 + q_3y_3^2 + \ldots + q_my_m^2, \\ c_1 = r_1z_1^2 + r_2z_2^2 + r_3z_3^2 + \ldots + r_nz_n^2, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \end{cases}$$

Now since the assumptions we have just made concerning the quantities $x, y, z, \ldots$ are the only ones possible, their probabilities will satisfy the following equations:

$$(3) \qquad \begin{cases} p_1 + p_2 + p_3 + \ldots + p_l = 1, \\ q_1 + q_2 + q_3 + \ldots + q_m = 1, \\ r_1 + r_2 + r_3 + \ldots + r_n = 1, \\ \ldots\ldots\ldots\ldots\ldots\ldots \end{cases}$$

It will now be easy for us to find by the aid of equations (1), (2), and (3), to what the sum of the values of the expression

$$(x_\lambda + y_\mu + z_\nu + \ldots - a - b - c - \ldots)^2 p_\lambda q_\mu r_\nu \ \ldots,$$

will reduce if we make successively

$$\lambda = 1, 2, 3, \ldots, l,$$
$$\mu = 1, 2, 3, \ldots, m,$$
$$\nu = 1, 2, 3, \ldots, n^1$$

Indeed when this expression is developed we have

$$\begin{aligned} &p_\lambda q_\mu r_\nu \ldots x_\lambda^2 + p_\lambda q_\mu r_\nu \ldots y_\mu^2 + p_\lambda q_\mu r_\nu \ldots z_\nu^2 + \ldots \\ &\quad + 2p_\lambda q_\mu r_\nu \ldots x_\lambda y_\mu + 2p_\lambda q_\mu r_\nu \ldots x_\lambda z_\nu + 2p_\lambda q_\mu r_\nu \ldots y_\mu z_\nu + \ldots \\ &\quad - 2(a + b + c + \ldots)p_\lambda q_\mu r_\nu \ldots x_\lambda \\ &\qquad\qquad\qquad\qquad - 2(a + b + c + \ldots)p_\lambda q_\mu r_\nu \ldots y_\mu \\ &\quad - 2(a + b + c + \ldots)p_\lambda q_\mu r_\nu \ldots z_\nu - \ldots \\ &\quad + (a + b + c + \ldots)^2 p_\lambda q_\mu r_\nu \ldots \end{aligned}$$

Giving to λ in this expression all the values from $\lambda = 1$ to $\lambda = l$, and summing the results of these substitutions, we will obtain the sum as follows:

[1] [The original has here $\nu + 1, 2, 3, \ldots, n \ldots$ which is obviously a misprint.]

$$
\begin{aligned}
q_\mu r_\nu \ldots & (p_1x_1^2 + p_2x_2^2 + p_3x_3^2 + \ldots + p_lx_l^2) \\
& + (p_1 + p_2 + p_3 + \ldots + p_l)q_\mu r_\nu \ldots y_\mu^2 \\
& + (p_1 + p_2 + p_3 + \ldots + p_l)q_\mu r_\nu \ldots z_\nu^2 \\
& + 2(p_1x_1 + p_2x_2 + p_3x_3 + \ldots + p_lx_l)q_\mu r_\nu \ldots y_\mu \\
& + 2(p_1x_1 + p_2x_2 + p_3x_3 + \ldots + p_lx_l)q_\mu r_\nu \ldots z_\nu \\
& + 2(p_1 + p_2 + p_3 + \ldots + p_l)q_\mu r_\nu \ldots y_\mu z_\nu \\
& + \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
& - 2(a + b + c + \ldots)(p_1x_1 + p_2x_2 + p_3x_3 + \ldots + p_lx_l) \\
& \qquad\qquad q_\mu r_\nu \ldots \\
& - 2(a + b + c + \ldots)(p_1 + p_2 + p_3 + \ldots + p_l)q_\mu r_\nu \ldots y_\mu \\
& - 2(a + b + c + \ldots)(p_1 + p_2 + p_3 + \ldots + p_l)q_\mu r_\nu \ldots z_\nu - \ldots \\
& + (a + b + c + \ldots)^2(p_1 + p_2 + p_3 + \ldots + p_l)q_\mu r_\nu \ldots
\end{aligned}
$$

If by means of equations (1), (2), and (3) we substitute in place of the sums

$$
\begin{gathered}
p_1x_1 + p_2x_2 + p_3x_3 + \ldots + p_lx_l, \\
p_1x_1^2 + p_2x_2^2 + p_3x_3^2 + \ldots + p_lx_l^2
\end{gathered}
$$

and

$$p_1 + p_2 + p_3 + \ldots + p_l$$

their values a, a_1 and 1, we will obtain the following formula:

$$
\begin{aligned}
a_1q_\mu r_\nu \ldots & + q_\mu r_\nu \ldots y_\mu^2 + q_\mu r_\nu \ldots z_\nu^2 + \ldots \\
& + 2aq_\mu r_\nu \ldots y_\mu + 2aq_\mu r_\nu \ldots z_\nu + 2q_\mu r_\nu \ldots y_\mu z_\nu + \ldots \\
& - 2(a + b + c \ldots)aq_\mu r_\nu \ldots - 2(a + b + c \ldots) \\
& \qquad\qquad q_\mu r_\nu \ldots z_\nu - \ldots \\
& + (a + b + c \ldots)^2 q_\mu r_\nu \ldots
\end{aligned}
$$

If we give to μ in this formula the values

$$\mu = 1, 2, 3, \ldots m,$$

then sum the expressions which result from these substitutions, and substitute for the sums

$$
\begin{gathered}
q_1y_1 + q_2y_2 + q_3y_3 + \ldots + q_my_m, \\
q_1y_1^2 + q_2y_2^2 + q_3y_3^2 + \ldots + q_my_m^2, \\
q_1 + q_2 + q_3 + \ldots + q_m,
\end{gathered}
$$

their values b, b_1, and 1 derived from equations (1), (2) and (3) we will obtain the following expression:

$$
\begin{aligned}
a_1r_\nu \ldots & + b_1r_\nu \ldots + r_\nu \ldots z_\nu^2 + \ldots \\
& + 2abr_\nu \ldots + 2ar_\nu \ldots z_\nu + 2br_\nu \ldots z_\nu + \ldots \\
& - 2(a + b + c + \ldots)ar_\nu \ldots - 2(a + b + c + \ldots)br_\nu \ldots \\
& - 2(a + b + c + \ldots)r_\nu \ldots z_\nu - \ldots + (a + b + c + \ldots)^2 r_\nu \\
& \qquad\qquad \ldots
\end{aligned}
$$

By treating ν in the same manner we will see that the sum of all the values of the expression

$$(x_\lambda + y_\mu + z_\nu + \ldots - a - b - c \ldots)^2 p_\lambda q_\mu r_\nu \ldots$$

derived by letting

$$\begin{aligned} \lambda &= 1, 2, 3, \ldots, l, \\ \mu &= 1, 2, 3, \ldots, m, \\ \nu &= 1, 2, 3, \ldots, n, \\ &\ldots\ldots \end{aligned}$$

will be equal to

$$\begin{aligned} a_1 + b_1 + c_1 + \ldots + 2ab + 2ac + 2bc + \ldots - 2(a + b + c \ldots)a \\ - 2(a + b + c \ldots)b - 2(a + b + c \ldots)c - \ldots \\ + (a + b + c \ldots)^2. \end{aligned}$$

Upon developing this expression it reduces to

$$a_1 + b_1 + c_1 + \ldots - a^2 - b^2 - c^2 \ldots$$

Hence we conclude that the sum of the values of the expression

$$\frac{(x_\lambda + y_\mu + z_\nu + \ldots - a - b - c - \ldots)^2}{\alpha^2(a_1 + b_1 + c_1 + \ldots - a^2 - b^2 - c^2 - \ldots)} p_\lambda q_\mu r_\nu \ldots,$$

which we obtain by making

$$\begin{aligned} \lambda &= 1, 2, 3, \ldots, l, \\ \mu &= 1, 2, 3, \ldots, m, \\ \nu &= 1, 2, 3, \ldots, n, \\ &\ldots\ldots \end{aligned}$$

will be equal to $\frac{1}{\alpha^2}\cdot$ Now it is evident that by rejecting from that sum all the terms in which the factor

$$\frac{(x_\lambda + y_\mu + z_\nu + \ldots - a - b - c - \ldots)^2}{\alpha^2(a_1 + b_1 + c_1 + \ldots - a^2 - b^2 - c^2 - \ldots)}$$

is less than 1 and by substituting unity for all those larger than 1, we will decrease that sum, and it will be less than $\frac{1}{\alpha^2}\cdot$ But this reduced sum will be formed of only those products $p_\lambda q_\mu r_\nu \ldots$, which correspond to the values of x_λ, y_μ, $z_\nu, \ldots$ for which the expression

$$\frac{(x_\lambda + y_\mu + z_\nu + \ldots - a - b - c - \ldots)^2}{\alpha^2(a_1 + b_1 + c_1 + \ldots - a^2 - b^2 - c^2 - \ldots)} > 1, \tag{4}$$

and it will evidently represent the probability that $x, y, z \ldots$ have values which satisfy condition (4).

This same probability can be replaced by $1 - P$, if we designate by P the probability that the values of $x, y, z, \ldots$ do not satisfy condition (4), or better—and this is the same thing—that the quantities have values for which the ratio

$$\frac{(x + y + z + \ldots - a - b - c - \ldots)^2}{\alpha^2(a_1 + b_1 + c_1 + \ldots - a^2 - b^2 - c^2 - \ldots)}$$

is not >1. Consequently the sum $x + y + z \ldots$ is included within the limits

$$a + b + c + \ldots + \alpha\sqrt{a_1 + b_1 + c_1 + \ldots - a^2 - b^2 - c^2 - \ldots}$$

and

$$a + b + c + \ldots - \alpha\sqrt{a_1 + b_1 + c_1 + \ldots - a^2 - b^2 - c^2 - \ldots}$$

Hence it is evident that the probability P must satisfy the inequality

$$1 - P < \frac{1}{\alpha^2},$$

which gives us

$$P > 1 - \frac{1}{\alpha^2},$$

which was to have been proved.

If N be the number of the quantities $x, y, z, \ldots$, and if in the theorem which we have just demonstrated we set

$$\alpha = \frac{\sqrt{N}}{t},$$

and divide by N both the sum $x + y + z + \ldots$ and its limits

$$a + b + c + \ldots + \alpha\sqrt{a_1 + b_1 + c_1 + \ldots - a^2 - b^2 - c^2 - \ldots}$$

and

$$a + b + c + \ldots - \alpha\sqrt{a_1 + b_1 + c_1 + \ldots - a^2 - b^2 - c^2 - \ldots},$$

we will obtain the following theorem concerning the mean values.

THEOREM. *If the mathematical expectations of the quantities $x, y, z, \ldots$ and $x^2, y^2, z^2, \ldots$ be respectively $a, b, c, \ldots a_1, b_1, c_1, \ldots$, be probability that the difference between the arithmetic mean of the N quantities $x, y, z, \ldots$ and the arithmetic mean of the mathematical expectations of these quantities will not exceed*

$$\frac{1}{t}\sqrt{\frac{a_1 + b_1 + c_1 + \ldots}{N} - \frac{a^2 + b^2 + c^2 + \ldots}{N}}$$

will always be larger than $1 - \frac{t^2}{N}$ whatever the value of t.

Since the fractions $\frac{a_1 + b_1 + c_1 + \ldots}{N}$ and $\frac{a^2 + b^2 + c^2 + \ldots}{N}$ express the mean of the quantities $a_1, b_1, c_1, \ldots$ and $a_1^2, b_1^2, c_1^2, \ldots$, whenever the mathematical expectations $a, b, c, \ldots a_1, b_1, c_1, \ldots$ do not exceed a given finite limit, the expression

$$\sqrt{\frac{a_1 + b_1 + c_1 + \ldots}{N} - \frac{a^2 + b^2 + c^2 + \ldots}{N}}$$

will have a finite value, no matter how large the number N, and in consequence we may make the value of

$$\frac{1}{t}\sqrt{\frac{a_1 + b_1 + c_1 + \ldots}{N} - \frac{a^2 + b^2 + c^2 + \ldots}{N}}$$

as small as we wish by giving to t a value sufficiently large. Now since, no matter what the value of t, if the number N approaches infinity the fraction $\frac{t^2}{N}$ will approach zero, by means of the preceding theorem, we reach the conclusion:

THEOREM. *If the mathematical expectations of the quantities* $U_1, U_2, U_3, \ldots$ *and of their squares* $U_1^2, U_2^2, U_3^2, \ldots$ *do not exceed a given finite limit, the probability that the difference between the arithmetic mean of* N *of these quantities and the arithmetic mean of their mathematical expectations will be less than a given quantity, becomes unity as* N *becomes infinite.*

For the particular hypothesis that the quantities U_1, U_2, $U_3, \ldots$ are either unity or zero, as when an event E either happens or fails on the 1st, 2nd, 3rd,...Nth *trial*, we note that the sum $U_1 + U_2 + U_3 + \ldots + U_N$ will give the number of *repetitions* of the event E in N trials, and that the arithmetic mean

$$\frac{U_1 + U_2 + U_3 + \ldots + U_N}{N}$$

will represent the ratio of the number of *repetitions* of the event E to the number of *trials*. In order to apply to this case our last theorem, let us designate by $P_1, P_2, P_3, \ldots P_N$ the probabilities of the event E in the 1st, 2nd, 3rd...Nth trial. The mathematical expectations of the quantities $U_1 + U_2 + U_3 + \ldots + U_N$ and of their squares $U_1^2, U_2^2, U_3^2, \ldots, U_N^2$ will be expressed, in conformity with our notation, as

$$P_1 1 + (1 - P_1)0,\ P_2 1 + (1 - P_2)0,\ P_3 1 + (1 - P_3)0,\ \ldots$$
$$P_1 1^2 + (1 - P_1)0^2, P_2 1^2 + (1 - P_2)0^2, P_3 1^2 + (1 - P_3)0^2, \ldots$$

Hence we see that the mathematical expectations are $P_1, P_2, P_3, \ldots$ and that the arithmetic mean of the first N expectations is

$$\frac{P_1 + P_2 + P_3 + \ldots + P_N}{N},$$

that is to say, the arithmetic mean of the probabilities P_1, P_2, $P_3, \ldots P_N$.

As a consequence of this, and by virtue of the preceding theorem, we arrive at the following conclusion:

When the number of trials becomes infinite, we obtain a probability —which may even be approximately one if we so wish—that the difference between the arithmetic mean of the probabilities of the event, during the trials, and the ratio of the number of repetitions of that event to the total number of trials, is less than any given quantity.

In the particular case in which the probability remains constant during all the trials, we have the theorem of Bernoulli.

LAPLACE

On the Probability of the Errors in the Mean Results of a Great Number of Observations and on the Most Advantageous Mean Result

(Translated from the French by Dr. Julian L. C. A. Gÿs, Harvard University, Cambridge, Mass.)

Pierre-Simon, Marquis de Laplace (1749–1827), born at Beaumont-en-Auge (Calvados), the son of a farmer, was in his early years professor of mathematics at the military school in his native city. He took part in the founding of the École Polytechnique and the École Normale. He dealt mostly with problems of celestial mechanics, and to the works of Newton, Halley, Clairaut, d'Alembert, and Euler on the consequences of universal gravitation, he added many personal contributions relating to the variations of the motion of the moon, the perturbations of the planets Jupiter and Saturn, the theory of the satellites of Jupiter, the velocity of the rotation of the ring of Saturn, aberration, the motion of the comets, and the tides. He was also famous for the invention of the cosmogonic system which bears his name. His *Théorie analytique des probabilités* ranks among the most important works done in the field of probability theory. In the edition of Mme. Vve. Courcier, Paris, 1820, it is preceded by a most interesting introduction which was first published under the title *Essai philosophique sur les probabilités.*
were published under the auspices of the Academy of Sciences in 1886.

The extract under consideration has been taken from the *Œuvres complètes de Laplace,* Vol. VII, published under the auspices of the Académie des Sciences, Paris, 1886 (Book 2, Chapter IV, pp. 304–327).

The interest in the passage lies in presenting the line of reasoning by which Laplace arrived at what is generally known as the law of errors of Gauss. Laplace certainly discovered the law before Gauss published his way of deriving it from his well-known postulates on errors. The method of Laplace is entirely different from that of Gauss. It should be noted that De Moivre gave a proof of the same law in 1733. (See page 566).

CHAPTER IV

On the probability of the errors in the mean results of a great number of observations and on the most advantageous mean result.

18. Let us now consider the mean results of a great number of observations of which the law of the frequency of the errors is known. Let us first assume that for each observation the errors may equally be

$$-n,\ -n+1,\ -n+2, \ldots -1,\ 0,\ 1, \ldots n-2,\ n-1,\ n.$$

The probability of each error will be $\frac{1}{2n+1}$. If we call the number of observations s, the coefficient of $c^{l\omega\sqrt{-1}}$ in the development of the polynomial

$$\left\{\begin{array}{l} c^{-n\omega\sqrt{-1}} + c^{-(n-1)\omega\sqrt{-1}} + c^{-(n-2)\omega\sqrt{-1}} \ldots \\ \ldots + c^{-\omega\sqrt{-1}} + 1 + c^{\omega\sqrt{-1}} \ldots + c^{n\omega\sqrt{-1}} \end{array}\right\}^s$$

will be the number of combinations in which the sum of the errors is l.[1] This coefficient is the term independent of $c^{\omega\sqrt{-1}}$ and of its powers in the development of the same polynomial multiplied by $c^{-l\omega\sqrt{-1}}$, and it is clearly equal to the term independent of ω in the same development multiplied by $\frac{c^{l\omega\sqrt{-1}} + c^{-l\omega\sqrt{-1}}}{2}$ or by $\cos l\omega$ Thus we have for the expression of this coefficient,

$$\frac{1}{\pi}\int d\omega.\cos l\omega.\,(1 + 2\cos\omega + 2\cos 2\,\omega \ldots + 2\cos n\omega)^s,$$

the integral being taken from $\omega = 0$ to $\omega = \pi$.

We have seen (Book I, art. 36) that this integral is[2]

$$\frac{(2n+1)^s\sqrt{3}}{\sqrt{n.(n+1).2s\pi}}.c^{-\frac{\frac{3}{2}l^2}{n(n+1)s}};$$

the total number of combinations of the errors is $(2n+1)^s$. Dividing the former quantity by the latter, we have for the probability that the sum of the errors of the s observations be l,

$$\frac{\sqrt{3}}{\sqrt{n(n+1).2s\pi}}.c^{-\frac{\frac{3}{2}l^2}{n(n+1).s}},$$

If we set

$$l = 2t.\sqrt{\frac{n.(n+1).s}{6}},$$

the probability that the sum of the errors will be within the limits $+2T\sqrt{\frac{n.(n+1).s}{6}}$ and $-2T.\sqrt{\frac{n.(n+1).s}{6}}$ will be equal to

$$\frac{2}{\sqrt{\pi}}.\int dt.c^{-t^2},$$

---

[1] [Here c stands for what we now represent by e.]

[2] [In section 36 of his Book I, Laplace computes the coefficient of $a^{\pm l}$ in the development of the polynomial

$$(a^{-n} + a^{-n+1} + \ldots + a^{-1} + 1 + a + \ldots + a^{n-1} + a^n)^s$$

where $a = c^{\omega\sqrt{-1}}$, in the case of a very large exponent.]

the integral being taken from $t = 0$ to $t = T$. This expression holds for the case of n infinite. Then, calling $2a$ the interval between the limits of the errors of each observation, we have $n = a$, and the preceding limits become $\pm \frac{2T.a.\sqrt{s}}{\sqrt{6}}$: thus the probability that the sum of the errors will be included within the limits $\pm ar.\sqrt{s}$ is

$$2.\sqrt{\frac{3}{2\pi}}.\int dr.c^{-\frac{3}{2}r^2}.$$

This is also the probability that the mean error will be included within the limits $\pm\frac{ar}{\sqrt{s}}$; for the mean error is obtained by dividing the sum of the errors by s.

The probability that the sum of the inclinations of the orbits of s comets will be included within the given limits, assuming that all inclinations are equally possible from zero to a right angle, is evidently the same as the preceding probability. The interval $2a$ of the limits of the errors of each observation is in this case the interval $\pi/2$ of the limits of the possible inclinations. Thus the probability that the sum of the inclinations will be included within the limits $\pm\frac{\pi.r\sqrt{s}}{4}$ is $2.\sqrt{\frac{3}{2\pi}}.\int dr.c^{-\frac{3}{2}r^2}$ which agrees with what has been found in section 13.[1]

Let us assume in general that the probability of each error positive or negative, may be expressed by $\phi(x/n)$, x and n being infinite numbers. Then, in the function

$$1 + 2\cos\omega + 2\cos 2\omega + 2\cos 3\omega \ldots + 2\cos n\omega,$$

each term such as $2\cos x\omega$ must be multiplied by $\phi(x/n)$. But we have

$$2\phi\left(\frac{x}{n}\right)\cdot\cos x\omega = 2\phi\left(\frac{x}{n}\right) - \frac{x^2}{n^2}\cdot\phi\left(\frac{x}{n}\right)\cdot n^2\omega^2 + \text{etc.}$$

Thus by putting

$$x' = \frac{x}{n}, \qquad dx' = \frac{1}{n},$$

[1] [In that section, Laplace finds the same result by considering the problem of the inclinations of the orbits as an application of this problem: given an urn containing $(n + 1)$ balls numbered from 0 to n, to find the probability of attaining the sum s by i drawings if the ball drawn is returned each time.]

the function

$$\phi\left(\frac{0}{n}\right) + 2\phi\left(\frac{1}{n}\right)\cdot \cos \omega + 2\phi\left(\frac{2}{n}\right)\cdot \cos 2\omega \ldots + 2\phi\left(\frac{n}{n}\right)\cdot \cos n\omega,$$

becomes

$$2n.\int dx'.\phi(x') - n^3\omega^2.\int x'^2dx'.\phi(x') + \text{etc.},$$

the integrals being taken from $x' = 0$ to $x' = 1$. Then let

$$k = 2\int dx'.\phi(x'), \qquad k'' = \int x'^2dx'.\phi(x'), \text{ etc.}$$

The preceding series becomes

$$nk.\left(1 - \frac{k''}{k}\cdot n^2\omega^2 + \text{etc.}\right).$$

Now the probability that the sum of the errors of s observations will lie within the limits $\pm l$ is, as is easily verified by the preceding reasoning,

$$\frac{2}{\pi}\cdot\int\int d\omega.dl. \cos l\omega \begin{Bmatrix} \phi\left(\frac{0}{n}\right) + 2\phi\left(\frac{1}{n}\right)\cdot \cos \omega + 2\phi\left(\frac{2}{n}\right)\cdot \cos 2\omega \ldots \\ \ldots + 2\phi\left(\frac{n}{n}\right)\cdot \cos n\omega \end{Bmatrix}^s$$

the integral being taken from $\omega = 0$ to $\omega = \pi$. This probability is then

$$2.\frac{(nk)^s}{\pi}\cdot\int\int d\omega.dl. \cos i\omega\left(1 - \frac{k''}{k}\cdot n^2\omega^2 - \text{etc.}\right)^s. \qquad (u)$$

Let us assume that

$$\left(1 - \frac{k''}{k}\cdot n^2\omega^2 - \text{etc.}\right)^s = c^{-t^2}.$$

In taking hyperbolic logarithms, when s is a large number, we have very nearly

$$s.\frac{k''}{k}\cdot n^2\omega^2 = t^2;$$

which yields

$$\omega = \frac{t}{n}\cdot\sqrt{\frac{k}{k''s}}.$$

If we then observe that the quantity nk or $2.\int dx.\phi(x/n)$ which expresses the probability that the error of an observation is

included within the limits $\pm n$, should be equal to unity, the function (u) becomes

$$\frac{2}{n\pi}\cdot\sqrt{\frac{k}{k''s}}\cdot\int\int dl.dt.c^{-t^2}.\cos\left(\frac{lt}{n}\cdot\sqrt{\frac{k}{k''s}}\right);$$

the integral with respect to t being taken between $t = 0$ and $t = \pi n\sqrt{\frac{k''s}{k}}$ or to $t = \infty$, n being assumed to be infinite. But from Book I, section 25[1] we have

$$\int dt.\cos\left(\frac{lt}{n}\cdot\sqrt{\frac{k}{k''s}}\right)\cdot c^{-t^2} = \frac{\sqrt{\pi}}{2}\cdot c^{-\frac{l^2}{4n^2}\cdot\frac{k}{k''s}}.$$

Then setting

$$\frac{l}{n} = 2t'.\sqrt{\frac{k''s}{k}};$$

the function (u) becomes

$$\frac{2}{\sqrt{\pi}}\cdot\int dt'.c^{-t'^2}.$$

Thus, calling the interval included between the limits of errors of each observation $2a$ as above, the probability that the sum of the errors of s observations will be included within the limits $\pm ar.\sqrt{s}$ is

$$\sqrt{\frac{k}{k''\pi}}\cdot\int dr.c^{-\frac{kr^2}{4k''}},$$

if $\phi\left(\frac{x}{n}\right)$ is constant. Then $\frac{k}{k''s} = 6$, and this probability becomes

$$2.\sqrt{\frac{3}{2\pi}}\int dr.c^{-\frac{3}{2}r^2},$$

which is conformable to what we found above.

If $\phi\left(\frac{x}{n}\right)$ or $\phi(x')$ is a rational and entire function of x', we have by the method of section 15, the probability that the sum of the errors shall be included within the limits $\pm ar.\sqrt{s}$ expressed by a series of powers of s, $2s$, etc., of quantities of the form

$$s - \mu \pm r.\sqrt{s},$$

[1] [Result found by Laplace in his chapter on integration by approximation of differentials which contain factors raised to high powers.]

in which μ increases in arithmetic progression, these quantities being continuous until they become negative. By comparing this series with the preceding expression of the same probability, we obtain the value of the series very accurately. And relative to this type of sequence we obtain theorems analogous to those which have been given in section 42, Book I, on the finite differences of powers of one variable.

If the law of frequency of the errors is expressed by a negative exponential that can extend to infinity and in general if the errors can extend to infinity then a becomes infinite and some difficulties may arise with the application of the preceding method. In all cases we shall set

$$\frac{x}{h} = x', \qquad \frac{1}{h} = dx',$$

h being an arbitrary finite quantity. And by following the above analysis exactly, we shall find for the probability that the sum of the errors of the s observations be included between the limits $\pm hr.\sqrt{s}$,

$$\sqrt{\frac{k}{k''\pi}} \cdot \int dr.c^{-\frac{kr^2}{4k''}},$$

an expression in which we should observe that $\phi\left(\frac{x}{h}\right)$ or $\phi(x')$ expresses the probability of the error $\pm x$, and that we have

$$k = 2\int dx'.\phi(x'), \qquad k'' = \int x'^2 dx'.\phi(x'),$$

the integrals being taken from $x' = 0$ to $x' = \infty$.

19. Let us now determine the probability that the sum of the errors of a very greater number of observations shall be included within the given limits, disregarding the signs of the errors, *i. e.*, taking them all as positive. To that end, let us consider the series

$$\phi\left(\frac{n}{n}\right).c^{-n\omega\sqrt{-1}} + \phi\left(\frac{n-1}{n}\right).c^{-(n-1)\omega\sqrt{-1}} \ldots + \phi\left(\frac{0}{n}\right)\ldots$$
$$\ldots + \phi\left(\frac{n-1}{n}\right).c^{(n-1)\omega\sqrt{-1}} + \phi\left(\frac{n}{n}\right).c^{n\omega\sqrt{-1}},$$

$\phi\left(\frac{x}{n}\right)$ being the ordinate of the probability curve of the errors, corresponding to the error $\pm x$, x being considered as well as n as formed by an infinite number of unities. By raising this series to the s-th power, after having changed the sign of the

negative exponentials, the coefficient of an arbitrary exponential, say $c^{(l+\mu s).\omega\sqrt{-1}}$, is the probability that the sum of the errors disregarding their sign, is $l + \mu s$; hence the probability is equal to

$$\frac{1}{2\pi}\cdot\int d\omega\cdot c^{-(l+\mu s)\omega\sqrt{-1}}\left\{\begin{array}{l}\phi\left(\frac{0}{n}\right)+2\phi\left(\frac{1}{n}\right)c^{\omega\sqrt{-1}}\\ +2\phi\left(\frac{2}{n}\right).c^{2\omega\sqrt{-1}}\ldots+2\phi\left(\frac{n}{n}\right).c^{n\omega\sqrt{-1}}\end{array}\right\}.$$

the integral with respect to ω being taken from $\omega = -\pi$ to $\omega = \pi$. Then, in that interval, the integral

$\int d\omega\cdot c^{-r\omega\sqrt{-1}}$, or $\int d\omega\cdot(\cos r\omega - \sqrt{-1}\sin r\omega)$, vanishes for all values of r that are not zero.

The development with respect to the powers of ω yields

$$\log\left\{c^{-\mu s\omega\sqrt{-1}}.\left[\phi\left(\frac{0}{n}\right)+2\phi\left(\frac{1}{n}\right)\cdot c^{\omega\sqrt{-1}}\ldots+2\phi\left(\frac{n}{n}\right)\cdot c^{n\omega\sqrt{-1}}\right]^s\right\}$$

$$= s\cdot\log\left\{\begin{array}{l}\phi\left(\frac{0}{n}\right)+2\phi\left(\frac{1}{n}\right)+2\phi\left(\frac{2}{n}\right)\ldots+2\phi\left(\frac{n}{n}\right)\\ +2\omega\sqrt{-1}\cdot\left[\phi\left(\frac{1}{n}\right)+2\phi\left(\frac{2}{n}\right)\ldots+n\phi\left(\frac{n}{n}\right)\right]\\ -\omega^2\cdot\left[\phi\left(\frac{1}{n}\right)+2^2\phi\left(\frac{2}{n}\right)\ldots+n^2\phi\left(\frac{n}{n}\right)\right]\end{array}\right\} - \mu s\omega\sqrt{-1}. \quad (1)$$

Therefore, setting $\frac{x}{n} = x'$, $\quad \frac{1}{n} = dx'$,

$2\int dx'\cdot\phi(x') = k$, $\quad \int x'dx'\cdot\phi(x') = k'$, $\quad \int x'^2dx'\cdot\phi(x') = k''$, $\int x'^3dx'\cdot\phi(x') = k'''$, $\int x'^4dx'\cdot\phi(x') = k^{IV}$, etc.

the integrals being taken from $x' = 0$ to $x' = 1$, the second member of the equation (1) becomes

$$s\cdot\log nk + s\cdot\log\left(1+\frac{2\cdot k'}{k}\cdot n\omega\sqrt{-1} - \frac{k''}{k}\,n^2\omega^2 - \text{etc.}\right) - \mu s\omega\sqrt{-1}.$$

As the error of each observation necessarily falls between the limits $\pm n$, we have $nk = 1$; and thus the preceding quantity becomes

$$s.\left(\frac{2k'}{k} - \frac{\mu}{n}\right).n\omega\sqrt{-1} - \frac{(kk'' - 2k'^2)\cdot s\cdot n^2\omega^2}{k^2} - \text{etc.}$$

By putting

$$\frac{\mu}{n} = \frac{2k'}{k},$$

and neglecting the powers of ω higher than the square, this quantity reduces to its second term and the preceding probability becomes

$$\frac{1}{2\pi}\int d\omega \cdot c^{-l\omega\sqrt{-1}\,-\,\frac{(kk''-2k'^2)}{k^2}\cdot s\cdot n^2\omega^2}$$

Let

$$\beta = \frac{k}{\sqrt{kk''-2k'^2}},\ \omega = \frac{\beta t}{n.\sqrt{s}},\qquad \frac{l}{n} = r.\sqrt{s}.$$

The preceding integral becomes

$$-\frac{\beta^2 r^2}{4} = \frac{1}{2\pi}\cdot\frac{c}{n.\sqrt{s}}\cdot\int \beta dt.c^{-\left(t+\frac{l\beta\sqrt{-1}}{2n\sqrt{s}}\right)^2}.$$

This integral should be taken from $t = -\infty$ to $t = \infty$; and then the preceding quantity becomes

$$\frac{\beta}{2\sqrt{\pi}.n.\sqrt{s}}\cdot c^{-\frac{\beta^2 r^2}{4}}.$$

On multiplying by dl or by $ndr.\sqrt{s}$ the integral

$$\frac{1}{2\sqrt{\pi}}\int \beta dr.c^{-\frac{\beta^2 r^2}{4}}$$

will be the probability that the value of l and consequently the sum of the errors of the observations is included between the limits $\frac{2k'}{k}\cdot as \pm ar.\sqrt{s}$, $\pm a$ being the limits of the errors of each observation, limits which we designate by $\pm n$ when we imagine them split up into an infinity of parts.

Thus we see that the most probable sum of the errors, disregarding the sign, is that which corresponds to $r = 0$. This sum is $\frac{2k'}{k}\cdot as$. In the case when $\phi(x)$ is constant, $\frac{2k'}{k} = \frac{1}{2}$, the most probable sum of the errors is then half of the largest possible sum, which sum is equal to sa. But if $\phi(x)$ is not a constant and if it decreases when the error x increases, then $\frac{2k'}{k}$ is less than $\frac{1}{2}$ and the sum of the errors disregarding the sign is less than half of the greatest sum possible.

By the same analysis, we can determine the probability that the sum of the squares of the errors will be $1 + \mu s$. It is easily seen that the expression of the probability is the integral

$$\frac{1}{2\pi}\cdot\int d\omega.c^{-(l+\mu s)\omega\sqrt{-1}}\left\{\begin{array}{r}\phi\left(\frac{0}{n}\right)+2\phi\left(\frac{1}{n}\right).c^{\omega\sqrt{-1}}+2\phi\left(\frac{2}{n}\right).c^{2^2\omega\sqrt{-1}}\\ \ldots+2\phi\left(\frac{n}{n}\right).c^{n^2\omega\sqrt{-1}}\end{array}\right\}^s,$$

taken from $\omega = -\pi$ to $\omega = \pi$. Following the preceding analysis precisely, we will have

$$\mu = \frac{2n^2.k''}{k};$$

and putting

$$\beta' = \frac{k}{\sqrt{kk^{IV} - 2k''^2}}$$

the probability that the sum of the squares of the errors of s observations will lie between the limits $\frac{2k''}{k}\cdot a^2s \pm a^2r.\sqrt{s}$ will be

$$\frac{1}{2\pi}\int \beta' dr.c^{-\frac{\beta'^2 r^2}{4}}.$$

The most probable sum is that which corresponds to $r = 0$ and therefore it is $\frac{2k''}{k}\cdot a^2s$. If s is an exceedingly large number, the result of the observations will differ very little from that value and will therefore yield very satisfactorily the factor $\frac{a^2\cdot k''}{k}\cdot$

20. When it is desired to correct an element already known to a good approximation by the totality of a great number of observations, we form equations of condition as follows: Let z be the correction of an element and let β be the observation, the analytic expression for which is a function of the element. By substituting for this element its approximate value plus the correction z and reducing to a series with respect to z and neglecting the square of z, this function will take the form $h + pz$. Setting it equal to the observed quantity β we obtain

$$\beta = h + pz;$$

z would thus be determined if the observation were exact, but since it is susceptible to error, we have exactly, calling that error ϵ to terms of order z^2

$$\beta + \epsilon = h + pz;$$

and by setting

$$\beta - h = \alpha,$$

we have

$$\epsilon = pz - \alpha.$$

Each observation yields a similar equation that may be written for the $(i + 1)$th observation as follows:

$$\epsilon^{(i)} = p^{(i)}\cdot z - \alpha^{(i)}.$$

Combining these equations, we have

$$S{\cdot}\epsilon^{(i)} = z{\cdot}S{\cdot}p^{(i)} - S{\cdot}\alpha^{(i)}, \tag{1}$$

where the symbol S holds for all values of i from $i = 0$ to $i = s - 1$, s being the total number of observations. Assuming that the sum of the errors is zero, this yields

$$z = \frac{S.\alpha^{(i)}}{S.p^{(i)}}.$$

This is what we usually call the *mean result of the observations.*

We have seen in section 18 that the probability that the sum of the errors of s observations be included within the limits $\pm ar.\sqrt{s}$ is

$$\sqrt{\frac{k}{k''\pi}}.\int dr.c^{-\frac{kr^2}{4k''}}.$$

Call $\pm u$ the error in the result z. Substituting $\pm ar.\sqrt{s}$ for $S.\epsilon^{(i)}$ in equation (1), and $\frac{S.\alpha^{(i)}}{S.p^{(i)}} \pm u$ for z, this yields

$$r = \frac{u.S.p^{(i)}}{a.\sqrt{s}}.$$

The probability that the error of the result z will be included within the limits $\pm u$ is thus

$$\sqrt{\frac{k}{k''s\pi}}.S.p^{(i)}.\int\frac{du}{a}.c^{-\frac{ku^2.(S.P^{(i)})^2}{4k''a^2s}}.$$

Instead of supposing that the sum of the errors is zero, we may suppose that an arbitrary linear function of these errors is zero, as

$$m.\epsilon + m^{(1)}.\epsilon^{(1)} + m^{(2)}.\epsilon^{(2)} \ldots + m^{(s-1)}.\epsilon^{(s-1)}, \tag{m}$$

$m, m^{(1)}, m^{(2)}$ being positive or negative integers. By substituting in this function (m) the values given by the equations of condition in the place of ϵ, $\epsilon^{(1)}$ etc., this becomes

$$z{\cdot}S{\cdot}m^{(i)}p^{(i)} - S{\cdot}m^{(i)}\alpha^{(i)}.$$

Setting the function (m) equal to zero yields

$$z = \frac{S{\cdot}m^{(i)}\alpha^{(i)}}{S{\cdot}m^{(i)}p^{(i)}}.$$

Let u be the error in this result so that

$$z = \frac{S.m^{(i)}\alpha^{(i)}}{S.m^{(i)}p^{(i)}} + u.$$

The function (m) becomes

$$u.S.m^{(i)}p^{(i)}.$$

Let us determine the probability of the error u, when the number of the observations is large.

For this, let us consider the product

$$\int\phi\left(\frac{x}{a}\right).c^{mx\omega\sqrt{-1}}\times\int\phi\left(\frac{x}{a}\right).c^{m^{(1)}x\omega\sqrt{-}}\ldots\times\int\phi\left(\frac{x}{a}\right).c^{m(s-1)x\omega\sqrt{-1}},$$

the symbol $\int$ extending over all the values of x, from the extreme negative value to the extreme positive value. As above, $\phi\left(\frac{x}{a}\right)$ is the probability of an error x in each observation, x being as is a, assumed to be formed of an infinite number of parts taken as unity. It is clear that the coefficient of an arbitrary exponential $c^{l\omega\sqrt{-1}}$ in the development of this product will be the probability that the sum of the errors of the observations, multiplied respectively by m, $m^{(1)}$ etc., in other words, the function (m), shall be equal to l. Then multiplying the latter product by $c^{-l\omega\sqrt{-1}}$, the term independent of $c^{\omega\sqrt{-1}}$ and of its powers in this new product will represent the same probability. If we assume as we did here, that the probability of positive errors is the same as that of negative errors, we may combine the terms multiplied by $c^{mx\omega\sqrt{-1}}$ and by $c^{-mx\omega\sqrt{-1}}$ in the sum $\int\phi\left(\frac{x}{a}\right).c^{mx\omega\sqrt{-1}}$. Then this sum will take the form $2\int\phi\left(\frac{x}{a}\right).\cos mx\omega$. And so for all similar sums. Hence the probability for the function (m) to be equal to l is

$$\frac{1}{2\pi}.\int d\omega\left\{\begin{array}{l} c^{-l\omega\sqrt{-1}}\times 2\int\phi\left(\frac{x}{a}\right).\cos mx\omega \\ \times 2\int\phi\left(\frac{x}{a}\right).\cos m^{(1)}x\omega\ldots\times 2\int\phi\left(\frac{x}{a}\right).\cos m^{(s-1)}x\omega \end{array}\right\};$$

the integral being taken from $\omega = -\pi$ to $\omega = \pi$. Reducing the cosines to a series yields

$$\int\phi\left(\frac{x}{a}\right).\cos mx\omega = \int\phi\left(\frac{x}{a}\right) - \frac{1}{2}.m^2a^2.\omega^2.\int\frac{x^2}{a^2}.\phi\left(\frac{x}{a}\right) + \text{etc.}$$

Letting $x/a = x'$ and observing that since the variation of x is unity, $dx' = 1/a$; we obtain

$$\int\phi\left(\frac{x}{a}\right) = a.\int dx'.\phi(x').$$

As above, let us call k the integral $2\int dx'.\phi(x')$ taken from $x' = 0$ to its extreme positive value, and k'' the integral $\int x'^2 dx'$ taken over the same limits, and so on, thus we will have

$$2\int \phi\left(\frac{x}{a}\right).\cos mx\omega = ak.\left(1 - \frac{k''}{k}.m^2a^2\omega^2 + \frac{k^{IV}}{12k}\cdot m^4a^4\omega^4 - \text{etc.}\right).$$

The logarithm of the second member of this equation is

$$-\frac{k''}{k}.m^2a^2\omega^2 + \frac{kk^{IV} - 6k''^2}{12k^2}.m^4a^4\omega^4 - \text{etc.} + \log ak,$$

ak or $2a\int dx'.\phi(x')$ expresses the probability that the error of each observation shall be included within the limits, a thing which is certain. We then have $ak = 1$. This reduces the preceding logarithm to

$$-\frac{k''}{k}.m^2a^2\omega^2 + \frac{kk^{IV} - 6k''^2}{12k^2}.m^4a^4\omega^4 - \text{etc.}$$

From this it is easy to conclude that the product

$$2\int \phi\left(\frac{x}{a}\right).\cos mx\omega \times 2\int \phi\left(\frac{x}{a}\right).\cos m^{(1)}x\omega \ldots \times 2\int \phi\left(\frac{x}{a}\right)\cos m^{(s-1)}x\omega,$$

is

$$\left(1 + \frac{kk^{IV} - 6k''^2}{12k^2}\cdot a^4\omega^4\cdot S\cdot m^{(i)4} + \text{etc.}\right).c^{-\frac{k''}{k}a^2\omega^2.S.m^{(i)2}}.$$

The preceding integral (i) reduces then to

$$\frac{1}{2\pi}\cdot\int d\omega.\left\{1 + \frac{kk^{IV} - 6k''^2}{12k^2}\cdot a^4\omega^4\cdot S\cdot m^{(i)4} + \text{etc.}\right\} \times c^{-l\omega\sqrt{-1} - \frac{k''}{k}.a^2\omega^2.S.m^{(i)2}}.$$

Setting $sa^2\omega^2 = t^2$, this integral becomes

$$\frac{1}{2a\pi\sqrt{s}}\int dt.\left\{1 + \frac{kk^{IV} - 6k''^2}{12k^2}\cdot\frac{S.m^{(i)4}}{s^2}.t^4 + \text{etc.}\right\} \times c^{-\frac{lt\sqrt{-1}}{a\sqrt{s}} - \frac{k''}{k}\cdot\frac{S.m^{(i)2}}{s}t^2};$$

$S.m^{(i)2}$, $S.m^{(i)4}$ are evidently quantities of order s. Thus $\frac{S.m^{(i)4}}{s^2}$ is of order $1/s$. Then neglecting the terms of the latter order with respect to unity, the above integral reduces to

$$\frac{1}{2a\pi.\sqrt{s}}\int dt.c^{-\frac{lt\sqrt{-1}}{a\sqrt{s}} - \frac{k''}{k}\cdot\frac{Sm^{(i)2}}{s}.t^2}.$$

The integral with respect to ω must be taken from $\omega = -\pi$ to $\omega = \pi$, the integral with respect to t must be taken from $t = -a\pi.\sqrt{s}$ to $t = a\pi.\sqrt{s}$, and in such cases the exponential under the radical sign is negligible at the two limits, either because s is a large number or because a is supposed to be divided up into an infinity of parts taken as unity. It is therefore permissible to take the integral from $t = -\infty$ to $t = \infty$. Letting

$$t' = \sqrt{\frac{k''.S.m^{(i)2}}{ks}}.\left\{t + \frac{l\cdot\sqrt{-1}\cdot k\cdot\sqrt{s}}{2a.k''\cdot S\cdot m^{(i)2}}\right\},$$

the preceding integral function becomes

$$\frac{c^{-\frac{kl^2}{4k''\cdot a^2\cdot S\cdot m^{(i)2}}}}{2a\pi\sqrt{\frac{k''}{k}.S.m^{(i)2}}}\cdot\int dt'.c^{-t'^2}.$$

The integral with respect to t' should be taken, just as the integral with respect to t, from $t' = -\infty$ to $t' = \infty$, so that the above quantity reduces to the following one

$$\frac{c^{-\frac{kl^2}{4k''.a^2.S.m^{(i)2}}}}{2a.\sqrt{\pi}\sqrt{\frac{k''}{k}.S.m^{(i)2}}}$$

Setting $l = ar.\sqrt{s}$ and observing that, since the variation of l is unity, $adr = 1$, we will have

$$\frac{\sqrt{s}}{2.\sqrt{\frac{k''\pi}{k}.S.m^{(i)2}}}\cdot\int dr.c^{-\frac{kr^2.s}{4k''.S.m^{(i)2}}}$$

for the probability that the function (m) be included within the limits zero and $ar.\sqrt{S}$, the integral being taken from r equal to zero.

Here we need to know the probability of the error u in the element as determined by setting the function (m) equal to zero. This function being assumed to be equal to l or to $ar.\sqrt{s}$, we have, according to previous relations

$$u.S.m^{(i)}p^{(i)} = ar.\sqrt{s}.$$

Substituting this value in the preceding integral function, this one becomes

$$\frac{S.m^{(i)}p^{(i)}}{2a\sqrt{\frac{k''\pi}{k}.S.m^{(i)2}}}\cdot\int du.c^{-\frac{ku^2.(S.m^{(i)}p^{(i)})^2}{4k''.a^2.S.m^{(i)2}}}.$$

This is the expression for the probability that the value of u be included between the limits zero and u. It is also the expression for the probability that u will be included between the limits zero and $-u$. Setting

$$u = 2at.\sqrt{\frac{k''}{k}}.\frac{\sqrt{S.m^{(i)2}}}{S.m^{(i)}p^{(i)}},$$

the preceding integral function becomes

$$\frac{1}{\sqrt{\pi}}\cdot\int dt.c^{-t^2}.$$

Now as the probability remains the same, t remains the same, and the interval of the two limits of u becomes smaller and smaller, the smaller $a.\sqrt{\frac{k''}{k}}.\frac{\sqrt{S.m^{(i)2}}}{S.m^{(i)}p^{(i)}}$ becomes. This interval remaining the same, the value of t and consequently the probability that the error of the element will fall within this interval, is the larger as the same quantity $a.\sqrt{\frac{k''}{k}}.\frac{\sqrt{S.m^{(i)2}}}{S.m^{(i)}p^{(i)}}$ is smaller. It is then necessary to chose a system of factors $m^{(i)}$ which will make this quantity a *minimum*. And as a, k, k'' are the same in all these systems, we must chose the system that will make $\frac{\sqrt{S.m^{(i)2}}}{S.m^{(i)}p^{(i)}}$ a minimum.

It is possible to arrive at the same result in the following way. Let us consider again the expression for the probability that u will be within the limits zero and u. The coefficient of du in the differential of that expression is the ordinate of the probability curve of the errors u in the element, errors which are represented by the abscissas of that curve which can be extended to infinity on both sides of the ordinate corresponding to $u = 0$. This being said, all errors whether positive or negative must be looked on as either a disadvantage or a real loss, in some game. Now by means of the probability theory, which has been expounded at some length in the beginning of this book, that disadvantage is computed by adding the products of each disadvantage by its

corresponding probability. The mean value of the error to fear in excess is thus equal to the integral

$$\frac{\int udu \cdot S \cdot m^{(i)}p^{(i)}.c^{-\frac{ku^2.(S.m^{(i)}p^{(i)})^2}{4k''.a^2.S.m^{(i)2}}}}{2a\sqrt{\frac{k''\pi}{k}.S.m^{(i)2}}}$$

the integral being taken from $u = 0$ to u infinite; thus the error is

$$a\sqrt{\frac{k''}{k\pi}}.\frac{\sqrt{S.m^{(i)2}}}{S.m^{(i)}p^{(i)}}.$$

The same quantity taken with the $-$ sign gives the mean error to fear in deficiency. It is evident that the system of the factors $m^{(i)}$ which must be chosen is such that these errors are *minima* and therefore such that

$$\frac{\sqrt{S.m^{(i)2}}}{S.m^{(i)}p^{(i)}}$$

is a *minimum*.

If we differentiate this function with respect to $m^{(i)}$ we will have, equating this derivative to zero, by the condition for a *minimum*,

$$\frac{m^{(i)}}{S.m^{(i)2}} = \frac{p^{(i)}}{S.m^{(i)}p^{(i)}}$$

This equation holds whatever i may be, and as the variation of i cannot affect the fraction $\frac{S.m^{(i)2}}{S.m^{(i)}p^{(i)}}$, setting this fraction equal to μ, we have

$$m = \mu \cdot p, \qquad m^{(1)} = \mu.p^{(1)}, \ldots \qquad m^{(s-1)} = \mu.p^{(s-1)};$$

and whatever p, $p^{(1)}$, etc. may be, we may take μ so that the numbers m, $m^{(1)}$, etc., are integers as the above analysis assumes. Then we have

$$z = \frac{S.p^{(i)}\alpha^{(i)}}{S.p^{(i)2}},$$

and the mean error to fear becomes $\pm \dfrac{a.\sqrt{\frac{k''}{k\pi}}}{\sqrt{S.p^{(i)2}}}.$ Under every hypothesis that can be made about the factors m, $m^{(i)}$, etc., this is the least mean error possible.

If we set the values of m, $m^{(1)}$, etc. equal to ± 1, the mean error to fear will be smaller when the sign $\pm$ is determined so that $m^{(i)}p^{(i)}$ will be positive. This amounts to supposing that

$1 = m = m^{(1)} =$ etc., and to preparing the equations of condition in such a way that the coefficient of z in each of them be positive. This is done by the ordinary method. The mean result of the observations is then

$$z = \frac{S.\alpha^{(i)}}{S.p^{(i)}},$$

and the mean error to fear whether it be in excess or in deficiency is equal to

$$\pm \frac{a.\sqrt{\frac{k''.s}{k\pi}}}{S.p^{(i)}}.$$

But this error is greater than the former which as has been seen is the smallest possible. Moreover this can be shown as follows. It suffices to prove the inequality

$$\frac{\sqrt{s}}{S.p^{(i)}} > \frac{1}{\sqrt{S.p^{(i)2}}},$$

or

$$s.S.p^{(i)2} > (S.p^{(i)})^2.$$

Indeed, $2pp^{(1)}$ is less than $p^2 + p^{(1)2}$ since $(p^{(1)} - p)^2$ is positive. Hence it is permissible to substitute for $2pp^{(1)}$ in the second member of the above inequality the quantity $p^2 + p^{(1)2} - f$, f being positive. Making similar substitutions for all similar products, that second member will be equal to the first one minus a positive quantity. The result

$$z = \frac{S.p^{(i)}\alpha^{(i)}}{S.p^{(i)2}}$$

to which corresponds the minimum of the mean error to fear, is the same as that given by the method of least squares of the errors of observations; for, as the sum of these squares is

$$(p\cdot z - \alpha)^2 + (p^{(1)}\cdot z - \alpha^{(1)})^2 \ldots + (p^{(s-1)}\cdot z - \alpha^{(s-1)})^2;$$

the *minimum* condition of this function yields when z varies, the preceding expression. Preference should thus be given to this method, for all laws of frequency of the errors whatsoever they may be are the laws on which the ratio $\frac{k''}{k}$ depends.

If $\phi(x)$ is a constant, this ratio is equal to $\frac{1}{6}$. It is less than $\frac{1}{6}$ if $\phi(x)$ varies in such a way that it decreases when x increases as

is natural to suppose. Adopting the mean law of errors given in section 15, according to which $\phi(x)$ is equal to $\frac{1}{2a}.\log\frac{a}{x}$, we have $\frac{k''}{k}=\frac{1}{18}$. As to the limits $\pm a$, we may take for those limits, the deviations from the mean result which would cause the rejection of an observation.

But, by means of the observations themselves, it is possible to determine the factor $a\cdot\sqrt{\frac{k''}{k}}$ in the expression for the mean error. Indeed it has been seen in the preceding section that the sum of the squares of the errors in the observations is very nearly equal to $2s.\frac{a^2k''}{k}$ and that it becomes extremely probable when there is a great number of observations for the observed sum not to differ from that value by an appreciable amount. We may set them equal to each other. Now the observed sum is equal to $S\cdot\epsilon^{(i)2}$ or to $S\cdot(p^{(i)}\cdot z-\alpha^{(i)})^2$. Substituting for z its value $\frac{S\cdot p^{(i)}\alpha^{(i)}}{S\cdot p^{(i)2}}$; it is found that

$$2s.\frac{a^2k''}{k}=\frac{S.p^{(i)2}.S.\alpha^{(i)2}-(S.p^{(i)}\alpha^{(i)})^2}{S.p^{(i)2}}.$$

The above expression for the mean error to fear in the result z then becomes

$$\pm\frac{\sqrt{S.p^{(i)2}.S.\alpha^{(i)2}-(S.p^{(i)}\alpha^{(i)})^2}}{S.p^{(i)2}.\sqrt{2s\pi}}$$

an expression in which nothing appears that is not given by the observations or by the coefficients of the equations of condition.

V. FIELD OF THE CALCULUS, FUNCTIONS, QUATERNIONS

Cavalieri's Approach to the Calculus

(Translated from the Latin by Professor Evelyn Walker, Hunter College, New York City.)

Bonaventura Francesco Cavalieri (Milan, 1598-Bologna, 1647), a Jesuit, was a pupil of Galileo. In order to prove his fitness for the Chair of Mathematics at the University of Bologna, he submitted, in 1629, the manuscript of his famous work, *Geometria Indivisibilibus Continuorum Nova quadam ratione promota*, which he published in 1635. This publication exerted an enormous influence upon the development of the calculus. Cavalieri was the author of a number of less important works, among them his *Exercitationes Geometricæ Sex*, which is still sometimes mentioned.

The following extract, known as Cavalieri's theorem, is from the *Geometria Indivisibilibus*, Book VII, Theorem 1, Proposition 1.[1]

Any plane figures, constructed between the same parallels, in which [plane figures] any straight lines whatever having been drawn equidistant from the same parallels, the included portions of any straight line are equal, will also be equal to one another; and any solid figures, constructed between the same parallel planes, in which [solid figures] any planes whatever having been drawn equidistant from the same parallel planes, the plane figures of any plane so drawn included within these solids, are equal, the [solid figures] will be equal to one another.

Now let the figures compared with one another, the plane as well as the solid, be called analogues, in fact even up[2] to the ruled lines or parallel planes between which they are assumed to lie, as it will be necessary to explain.

---

[1] This translation has been compared with and checked by that of G. W. Evans in *The American Mathematical Monthly*, XXIV, 10 (December 1917), pp. 447–451. The diagram and lettering used by Evans have been adopted as being more convenient than those of Cavalieri, whose diagram is not only poorly printed, but has its points designated by numbers as well as by both Roman and Greek letters.

[2] The expression used is "juxta regulas lineas," literally "next to the ruled lines."]

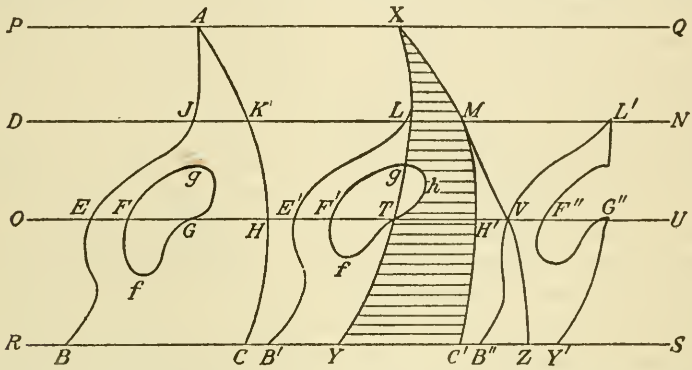
Let there be any plane figures, ABC, XYZ, constructed between the same parallels, PQ, RS; but DN, OU, any parallels to PQ, RS, having been drawn, the portion [s], for example of the DN, included within the figures, namely JK, LM, are equal to each other, and besides, the portions EF, GH, of the OU taken together (for the figure ABC, for example, may be hollow within following the contour FgG), are likewise equal to the TV; and let this happen in any other lines equidistant from the PQ. I say that the figures ABC, XYZ, are equal to each other

For either of the figures ABC, XYZ, as the ABC, having been taken with the portions of the parallels PQ, RS, coterminous with it, namely with PA, RB, let it be superimposed upon the remaining figure XYZ, but so that the [lines] PA, RB, may fall upon AQ and CS; then either the whole [figure] ABC coincides with the whole [figure] XYZ, and so, since they coincide with each other, they are equal, or not; yet there may be some part which coincides with another part, as $XMC'YTbL$, a part of the figure ABC, with $XMC'YTbL$, a part of the figure XYZ.

By the superposition of the figures effected in such a way that portions of the parallels PQ, RS, coterminous with the two figures, are superposed in turn, it is evident that whatever straight lines included within the figures were in line with each other, they still remain in line with each other, as, for example, since EF, GH, were in the same line TV, the said superposition having been made, they will still remain in line with themselves, obviously $E'F'$, TH', in line with the TV, for the distance of the EF, GH, from PQ is equal to the distance [of] TV from the same PQ. Whence, however many times PA is laid upon AQ, wherever it may be done, EF, GH, will always remain in line with the TV; which is clearly apparent also for any other lines whatsoever parallel to PQ in each figure.

But when a part of one figure, as ABC, necessarily coincides with a part of the figure XYZ and not with the whole, while the superposition is made according to such a rule as has been stated, it will be demonstrated thus. For since, any parallels whatever having been drawn to the [line] AD, the portions of them included within the figures, which were in line with one another still remain in line with one another after the superposition, they, of course, being equal by hypothesis before superposition, then after superposition the portions of the [lines] parallel to the AD, included within the superposed figures, will likewise be equal; as, for

example, $E'F'$, TH', taken together, will be equal to the TV; therefore if the $E'F'$, TH', together do not coincide with the whole TV, then a part coinciding with some part, as TH' with TH' itself, $E'F'$ will be equal to the $H'V$, and in fact $E'F'$ will be in the residuum of the superposed figure ABC, [and] $H'V$ indeed in the residuum of the figure XYZ upon which the superposition was made. In the same way we shall show [that] to any [line] whatsoever parallel to the PQ, included within the residuum of the superposed figure ABC, as it were $LB'YTF'$, there corresponds an equal straight line, in line [with it], which will be in the residuum of the figure XYZ upon which the superposition was made; therefore, the superposition having been carried out in accordance with this rule, when there is left over any [part] of the superposed figure



which does not fall on the figure upon which the superposition was made, it must be that some [part] of the remaining figure also is left over, upon which nothing has been superposed.

Since, moreover, to each one of the straight lines parallel to PQ included within the residuum or residuua (because there may be several residual figures) of the superposed figure ABC or XYZ, there corresponds, in line [with it], another straight line in the residuum or residua of the figure XYZ, it is manifest that these residual figures, or aggregates of residua, are between the same parallels. Therefore since the residual figure $LB'YTF'$ is between the parallels DN, RS, likewise the residual figure or aggregate of the residual figures of the XYZ, because it has within it the frusta Thg, $MC'Z$, will be between the same parallels DN, RS. For if it did not extend both ways to the parallels DN, RS, as, for example, if it extended indeed all the way to DN, but not all the way to RS, but only as far as OU, to the straight lines included

within the frustum $E'B'YfF'$, parallel to the [line] PQ, there would not correspond in the residuum of the figure XYZ, or [in] the aggregate of the residua, other straight lines, as, it was proved above, is necessary. Therefore these residua or the aggregates of residua are between the same parallels, and the portions of the [lines] included therein parallel to the PQ, RS, are equal to one another, as we have shown above. Therefore the remainders or the aggregates of the remainders, are in that condition in which, it was assumed just now, were the figures ABC and XYZ; that is, [they are] likewise analogues.

Then again let the superposition of the residua be made, but so that the parallels KL, CY, may be placed upon the parallels LN, YS, and the part $VB''Z$ of the frustum $LB'YTF'$ may coincide with the part $VB''Z$ of the frustum $MC'Z$. Then we shall show, as above, [that] while there is a residuum of one there is also a residuum of the other, and these residua, or aggregates of residua, are between the same parallels. Now let $L'VZY'G''F''$ be the residuum with respect to the figure ABC, but let the residua $MC'B''V$, Tbg, whose aggregate is between the same parallels as the residuum $L'VZY'G''F''$, belong to the figure XYZ, of course between the parallels DN, RS; then if again a superposition of these residua is made, but so that the parallels between which they lie may always be superposed in turn, and it may be understood that this is always to be done, until the whole figure ABC will have been superposed, I say [that] the whole [of it] must coincide with the XYZ; otherwise, if there is any residuum, as of the figure XYZ, upon which nothing had been superposed, then there would be some residuum of the figure ABC which would not have been superposed, as we have shown above to be necessary. But it has been stated that the whole ABC was superposed upon the XYZ; therefore they are so superposed one [part] after another, that there is a residuum of neither; therefore they are so superposed that they coincide with each other; therefore the figures ABC, XYZ, are equal to each other.

Now let there be constructed in the same diagram any two solid figures whatever, ABC, XYZ, between the same parallel planes PQ, RS; then any planes DN, OU, having been drawn equidistant from the aforesaid [planes], let the figures which are included within the solids, and which lie in the same plane, always be equal to each other, as JK equal to LM, and EF, GH, taken together (for a solid figure, for example ABC, may be hollow in any way

within, following the surface $FfGg$) equal to the [figure] TV. I say that these solid figures are equal.

For if we superpose the solid ABC, together with the portions PA, RC of the planes PQ, RS, coterminous with it, upon the solid XYZ, so that the plane PA may be upon [the plane] PQ, and [the plane] RC on the plane RS, we shall show (as we did above with respect to the portions of the lines parallel to the PQ included within the plane figures ABC, XYZ) that the figures included within the solids ABC, XYZ, which were in the same plane, will also remain in the same plane after superposition, and therefore thus far the figures included within the superposed solids and parallel to the PQ, RS, are equal.

Then unless the whole solid coincides with the [other] whole [solid] in the first superposition, there will remain residual solids, or solids composed of residua in either solid, which will not be superposed upon one another; for when, for example, the figures $E'F'$, TH', are equal to the figure TV, the common figure TH' having been subtracted, the remainder $E'F'$ will be equal to the remainder $H'V$; and this will happen in any plane whatsoever parallel to the plane PQ, meeting the solids ABC, XYZ. Therefore having a residuum of one solid, we shall always have a residuum of the other also. And it will be evident, according to the method applied in the first part of this proposition concerning plane figures, that the residual solids or the aggregates of the residua will always be between the same parallel planes, as the residua $LB'YTF'$, $MC'Z$, Thg, are between the parallel planes DN, RS, and likewise [that they are] analogues.

If therefore these residua also are superposed so that the plane DL is placed upon the plane LN, and RY upon YS, and this is understood to be done continually until [the one] which is superposed, as ABC, taken all together, as the whole ABC, will be coincident with the whole XYZ. For the whole solid ABC having been superposed upon the XYZ, unless they coincide with one another, there will be some residuum of one, as of the solid XYZ, and therefore there would be some residuum of the solid $XB'C'$ or ABC, and this would not have been superposed, which is absurd; for it has already been stated that the whole solid ABC was superposed upon the XYZ. Therefore there will not be any residuum in these solids. Therefore they will coincide. Therefore the said solid figures ABC, XYZ, will be equal to each other. Which [things] were to be demonstrated.

FERMAT

On Maxima and Minima

(Translated from the French by Dr. Vera Sanford, Teachers College, Columbia University, New York City.)

Supplementing the communication from Fermat to Pascal (see page 289), giving some of his ideas on analytic geometry, the following letter to Roberval shows how his mind was working toward one of the combinations of the calculus with the Cartesian system. The letter was written on Monday, September 22, 1636, a year before Descartes published *La Geometrie*. See the *Œuvres de Fermat* (ed. Tannery and Henry, Vol. II, pp. 71–74, Paris, 1894). For a biographical sketch of Fermat, see page 397; for the introductory pages of Descartes, see pages 213, 214.

Monsieur,

1. With your permission, I shall postpone writing you on the subject of the propositions of mechanics until you shall do me the favor of sending me the demonstration of your theorems which I trust to see as soon as possible according to the promise you made me.

2. On the subject of the method of *maxima* and *minima*, you know that as you have seen the work which M. Despagnet gave you, you have seen mine which I sent him about seven years since at Bordeaux.

At that time, I recollect that M. Philon received a letter from you in which you proposed that he find the greatest cone of all those whose conical surface is equal to a given circle. He forwarded it to me and I gave the solution to M. Prades to return to you. If you search your memory, you will perhaps recall it and also the fact that you set this question as a difficult one that had not then been solved. If I discover your letter, which I kept at the time, among my papers, I will send it to you.

3. If M. Despargnet laid my method before you as I then sent it to him, you have not seen its most beautiful applications, for I have made use of it by amplifying it a little:

(1) For the solution of problems such as that of the conoid which I sent you in my last letter.

(2) For the construction of tangents to curved lines on which subject I set you the problem: To draw a tangent at a given point on the conchoid of Nicomedes.

(3) For the discovery of the centers of gravity of figures of every type, even of figures that differ from the ordinary ones such as my conoid and other infinite figures of which I shall show examples if you desire them.

(4) For numerical problems in which there is a question of aliquot parts and which are all very difficult.

4. It is by this method that I discovered 672 [the sum of] whose factors are twice the number itself, just as the factors of 120 are twice 120.

It is by the same method also that I discovered the infinite numbers that make the same thing[1] as 220 and 284, that is to say, the [sum of the] factors of the first are equal to the second, and those of the second are equal to the first. If you wish to see an example of this to test the question, these two numbers 17296 and 18416 satisfy the conditions.

I am certain that you told me that this question and others of its type are very difficult. I sent the solution to M. de Beaugrand some time ago.

I have also found numbers which exceeded the aliquot parts of a given number in a given ratio, and several others.

5. These are the four types of problems included in my method which perhaps you did not know.

With reference to the first, I have squared[2] infinite figures bounded by curved lines; as for example, if you imagine a figure such as a parabola, of such a type that the cubes of the ordinates[3] are in proportion to the segments which they cut from the diameter. This figure is something like a parabola and it differs from one only in the fact that in a parabola we take the ratio of the squares while in this figure I take that of the cubes. It is for this reason that M. de Beaugrand, to whom I put this problem calls it a *cubical parabola*.[4]

I have also proved that this figure is in the sesquialter[5] ratio to the triangle of the same base and height. You will find on investi-

[1] [The "amicable numbers" 220 and 284 were known at an early date, possibly as early as the Pythagoreans. The second pair were discovered by Fermat whose expression "the infinite numbers" meant that the author had discovered a general rule for the formation of these numbers.]

[2] [Found the area of.]

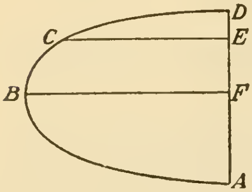
[3] [. . .que les cubes des appliquées soient en proportion des lignes qu'elles coupent du diamètre."]

[4] ["parabole solide."]

[5] [*I. e.*, the ratio of 3:2.]

gation that it was necessary for me to follow another method than that of Archimedes for the quadrature of the parabola, and that I should never have found it by his method.

6. Since you found my theorem on the conoid excellent, here is its most general case: If a parabola with the vertex B and axis BF and ordinate AF be revolved about the straight line AD, a new type of conoid will be produced in which a section cut by a plane perpendicular to the axis will have the ratio to the cone on the same base and with the same altitude that 8 has to 5.



If, indeed, the plane cuts the axis in unequal segments, as at E, the segment of the conoid $ABCE$ is to the cone of the same base and altitude as five times the square ED with twice the rectangle AED and the rectangle of DF and AE is to five times the square on ED. And likewise the segment of the conoid DCE is to the cone with the same base and altitude as five times the square AE added to twice the rectangle AED and the rectangle of DF and DE is to five times the square on AE.

For the proof, besides the aid which I have from my method, I make use of inscribed and circumscribed cylindars.

7. I have omitted the principal use of my method which is in the discovery of plane and solid loci. It had been of particular service to me in finding the plane loci which I found so difficult before:

If from any number of given points, straight lines are drawn to a single point, and if these lines are such that they are equally spaced by a given amount from each other, then the point lies on a circle which is given in position.

All that I shall tell you are but examples, for I can assure you that for each of the preceding points I have found a very great number of exceedingly beautiful theorems. I shall send you their proof if you wish. May I however beg you to try them soon and to give me your solutions.

8. Finally, since the last letter which I wrote you, I have discovered the theorem which I set you. It caused me the greatest difficulty and it did not occur to me at an earlier date.

I beg you to share some of your reflections with me and to believe me etc.

(Facing page 613.)

NEWTON

On Fluxions

(Translated from the Latin by Professor Evelyn Walker, Hunter College, New York City.)

Sir Isaac Newton (1642–1727) was the son of a Lincolnshire farmer. In 1660 he entered Trinity College, Cambridge, where he became the pupil of Isaac Barrow by whom his future work was strongly influenced. His discoveries in mathematics and physics began as early as 1664, although he did not publish any of his work until many years later. In 1669 he succeeded Barrow as Lucasian professor of mathematics at Trinity. Later he became warden of the mint and member of parliament, and was knighted by Queen Anne. He was elected fellow of the Royal Society in 1672, and from 1703 until his death was its president. In 1699 he was made foreign associate of the Académie des Sciences. He is buried in Westminister Abbey.[1]

His best known work is the *Principia*, or to give it its full title, *Philosophiæ Naturalis Principia Mathematica*, published in 1687, containing his theory of the universe based on his law of gravitation. Every high-school boy knows his name in connection with the binomial theorem, and more advanced students in connection with infinite series and the theory of equations. But his mathematical fame is due most of all to his invention of the calculus. His first development of the subject proceeded by means of infinite series as told by Wallis in his *Algebra*, 1685, Later he used the method that is most commonly associated with his name, that of fluxions, as exemplified in the *Quadratura Curvarum*, 1704.[2] Both of these naturally entail the use of infinitely small quantities. Finally in his *Principia* he explains the use of prime and ultimate ratios. The following quotations from the sources specified show the three points of view.

[*Integration by Means of Infinite Series*[3]]

He doth therein, not only give us many such Approximations... but he lays down general Rules and Methods... And gives

[1] For a brief summary of the life of Newton and a good bibliography for the same, see David Eugene Smith, *History of Mathematics*, Vol. I, p. 398.

[2] Charles Hayes published this method of Newton in a work of his own, *A Treatise of Fluxions: or, An Introduction to Mathematical Philosophy*, London, 1704. Nine years after Newton's death John Colson published *The Method of Fluxions and Infinite Series...from the Author's Latin Original not yet made publick...*, London, 1736.

[3] [John Wallis (see page 46) says in his *Algebra* (1685, p. 330) that he had seen the two letters written by Newton to Oldenburg, June 13 and October 24, 1676, containing Newton's discoveries in the realm of infinite series. The quotation is from Wallis.]

instances, how those Infinite or Interminate Progressions may be accommodated, to the Rectifying of Curve Lines...; Squaring of Curve-lined Figures; finding the Length of Archs,....

. . . .

[*Newton's Method of Fluxions*[1]]

Therefore, considering that quantities, which increase in equal times, and by increasing are generated, become greater or less according to the greater or less velocity with which they increase and are generated; I sought a method of determining quantities from the velocities of the motions, or [of the] increments, with which they are generated; and calling these velocities of the motions, or [of the] increments, *fluxions*, and the generated quantities *fluents*, I fell by degrees, in the years 1665 and 1666, upon the method of fluxions, which I have made use of here in the quadrature of curves.

Fluxions are very nearly as the augments of the fluents generated in equal, but very small, particles of time; and, to speak accurately, they are in the *first ratio* of the nascent augments; but they may be expounded by any lines which are proportional to them.

. . . .

It amounts to the same thing if the fluxions are taken in the ultimate ratio of the evanescent parts.[2]

. . . .

Let the quantity x flow uniformly, and let it be proposed to find the fluxion of x^n.

In the time that the quantity x, by flowing, becomes $x + o$, the quantity x^n will become $\overline{x + o}|^n$, that is, by the method of infinite series,

$$x^n + nox^{n-1} + \frac{n^2 - n}{2}oox^{n-2} + \text{etc.}$$

[1] [The quotation that follows is from *Quadratura Curvarum*, published with Newton's *Opticks: or, a Treatise of the Reflexions, Refractions, Inflexions and Colours of Light. Also Two Treatises of the Species and Magnitude of Curvilinear Figures*, London, 1704. The second of the treatises mentioned is the *Quadratura Curvarum*, pp. 165–211. We quote from the Introduction and give one proposition from the work itself. The translation as given here does not differ, except in a few unimportant details, from that of John Stewart, published in London, 1745.]

[2] [Newton here gives some examples. He makes the tangent coincide with the limiting position of the secant by making the ordinate of the second point of intersection of the secant with the curve move up into coincidence with that of the first. See page 617.]

And the augments o and $nox^{n-1} + \frac{n^2 - n}{2} oox^{n-2} +$ etc. are to one another as

$$1 \text{ and } nx^{n-1} + \frac{n^2 - n}{2} ox^{n-2} + \text{etc.}$$

Now let these augments vanish, and their ultimate ratio will be as

$$1 \text{ to } nx^{n-1}.$$

From the fluxions to find the fluents is a much more difficult problem, and the first step of the solution is to find the quadrature of curves; concerning which I wrote the following some time ago.[1]

In what follows I consider indeterminate quantities as increasing or decreasing by a continued motion, that is, by flowing or ebbing, and I designate them by the letters z, y, x, v, and their fluxions or velocities of increasing I denote by the same letters pointed $\dot{z}$, $\dot{y}$, $\dot{x}$, $\dot{v}$. There are likewise fluxions or mutations, more or less swift, of these fluxions, which may be called the second fluxions of the same quantities z, y, x, v, and may be thus designated: $\ddot{z}$, $\ddot{y}$, $\ddot{x}$, $\ddot{v}$; and the first fluxions of these last, or the third fluxions of z, y, x, v, thus: $\dddot{z}$, $\dddot{y}$, $\dddot{x}$, $\dddot{v}$; and the fourth fluxions thus: $\ddddot{z}$, $\ddddot{y}$, $\ddddot{x}$, $\ddddot{v}$. And after the same manner that $\dddot{z}$, $\dddot{y}$, $\dddot{x}$, $\dddot{v}$ are the fluxions of the quantities $\ddot{z}$, $\ddot{y}$, $\ddot{x}$, $\ddot{v}$, and these the fluxions of the quantities $\dot{z}$, $\dot{y}$, $\dot{x}$, $\dot{v}$, and these last the fluxions of the quantities z, y, x, v; so the quantities z, y, x, v, may be considered as the fluxions of others which I shall designate thus: $\acute{z}$, $\acute{y}$, $\acute{x}$, $\acute{v}$; and these as fluxions of others $\acute{\acute{z}}$, $\acute{\acute{y}}$, $\acute{\acute{x}}$, $\acute{\acute{v}}$; and these last as the fluxions of still others $\acute{\acute{\acute{z}}}$, $\acute{\acute{\acute{y}}}$, $\acute{\acute{\acute{x}}}$, $\acute{\acute{\acute{v}}}$. Therefore $\acute{\acute{z}}$, $\acute{z}$, z, $\dot{z}$, $\ddot{z}$, $\dddot{z}$, $\ddddot{z}$, $\overset{:\cdot:}{z}$ etc., designate a series of quantities whereof every one that follows is the fluxion of the preceding, and every one that goes before is a flowing quantity having the succeeding one as its fluxion.

. . . .

And it is to be remembered that any preceding quantity in this series is as the area of a curvilinear figure of which the succeeding quantity is the rectangular ordinate, and [of which] the abscissa is z;. . .

[1] [This is the end of the introduction.]

Prop. 1. *Prob.* 1. *An equation being given involving any number of flowing quantities, to find the fluxions.*

Solution. Let every term of the equation be multiplied by the index of the power of every flowing quantity that it involves, and in every multiplication let a side [or root] of the power be changed into its fluxion, and the aggregate of all the products, with their proper signs, will be the new equation.

Explication. Let a, b, c, d, etc., be determinate and invariable quantities, and let any equation be proposed involving the flowing quantities z, y, x, v, etc., as

$$x^3 - xy^2 + a^2z - b^3 = 0.$$

Let the terms be first multiplied by the indices of the powers of x, and in every multiplication, for the root or x of one dimension, write $\dot{x}$, and the sum of the terms will be

$$3\dot{x}x^2 - \dot{x}y^2.$$

Let the same be done in y, and it will produce

$$-2xy\dot{y}.$$

Let the same be done in z, and it will produce

$$aa\dot{z}.$$

Let the sum of these results be placed equal to nothing, and the equation will be obtained

$$3\dot{x}x^2 - \dot{x}y^2 - 2xy\dot{y} + aa\dot{z} = 0.$$

I say that the relation of the fluxions is defined by this equation.

Demonstration.—For let o be a very small quantity, and let $o\dot{z}$, $o\dot{y}$, $o\dot{x}$, be the moments, that is the momentaneous synchronal increments, of the quantities z, y, x. And if the flowing quantities are just now z, y, x, these having been increased after a moment of time by their increments $o\dot{z}$, $o\dot{y}$, $o\dot{x}$, these quantities will become $z + \dot{z}o$, $y + o\dot{y}$, $x + o\dot{x}$; which being written in the first equation for z, y and x, give this equation:

$$x^3 + 3x^2o\dot{x} + 3xoo\dot{x}\dot{x} + o^3\dot{x}^3 - xy^2 - o\dot{x}y^2 - 2xo\dot{y}y - 2xo^2\dot{y}y$$
$$- \dot{x}o^2\dot{y}\dot{y} - \dot{x}o^3\dot{y}\dot{y} + a^2z + a^2o\dot{z} - b^3 = 0.$$

Let the former equation be subtracted [from the latter] and the remainder be divided by 0, and it will be

$$3\dot{x}x^2 + 3\dot{x}\dot{x}ox + \dot{x}^3o^2 - \dot{x}y^2 - 2x\dot{y}y - 2\dot{x}o\dot{y}y - xo\dot{y}\dot{y} - \dot{x}o^2\dot{y}\dot{y}$$
$$+ a^2z = 0.$$

Let the quantity o be diminished infinitely, and, neglecting the terms which vanish, there will remain

$$3\dot{x}x^2 - \dot{x}y^2 - 2x\dot{y}y + a^2\dot{z} = 0.$$

Q. E. D.[1]

If the points are distant from each other by an interval, however small, the secant will be distant from the tangent by a small interval. That it may coincide with the tangent and the last ratio be found, the two points must unite and coincide altogether. In mathematics errors, however small, must not be neglected.

. . . .

[*The Method of Prime and Ultimate Ratios*[2]]

Quantities, as also ratios of quantities, which constantly tend toward equality in any finite time, and before the end of that time approach each other more nearly than [with] any given difference whatever, become ultimately equal...

The objection is that there is no ratio[3] of evanescent quantities, which obviously, before they have vanished, is not ultimate; when they have vanished, there is none. But also by the same like argument it may be contended that there is no ultimate velocity of a body arriving at a certain position; for before the body attains the position, this is not ultimate; when it has attained [it], there is none. And the answer is easy: By ultimate velocity I understand that with which the body is moved, neither before it arrives at the ultimate position and the motion ceases, nor thereafter, but just when it arrives; that is, that very velocity with which the body arrives at the ultimate position and with which the motion ceases. And similarly for the motion of evanescent quantities is to be understood the ratio of the quantities, not before they vanish, nor thereafter, but [that] with which they vanish. And likewise the first nascent ratio is the ratio with which they begin. And the prime and ultimate amount is to be [that] with which they begin and cease (if you will, to increase and diminish). There

---

[1] [The *Quadratura Curvarum*, 1704, besides explaining the method of fluxions, also anticipates the method of prime and ultimate ratios, which is practically the modern method of limits. The paragraph which follows occurs earlier in the text.]

[2] [This translation has been made from *Philosophiæ Naturalis Principia Mathematica Auctore Isaaco Newtono*. Amsterdam, 1714. The first edition was published in 1687. The selections are from pages 24 and 33.]

[3] [Newton's word is "proportio."]

exists a limit which the velocity may attain at the end of the motion, but [which it may] not pass. This is the ultimate velocity. And the ratio of the limit of all quantities and proportions, beginning and ceasing, is equal...

The ultimate ratios in which quantities vanish, are not really the ratios of ultimate quantities, but the limits toward which the ratios of quantities, decreasing without limit, always approach; and to which they can come nearer than any given difference, but which they can never pass nor attain before the quantities are diminished indefinitely.[1]

---

[1] Acknowledgment is hereby made to G. H. Graves, whose article, "Development of the Fundamental Ideas of the Differential Calculus," in *The Mathematics Teacher*, Vol. III (1910–1911), pp. 82–89, has been freely used.

Leibniz.

(Facing page 619.)

LEIBNIZ

On the Calculus

(Translated from the Latin by Professor Evelyn Walker, Hunter College, New York City.)

Gottfried Wilhelm, Freiherr von Leibniz (Leipzig, 1646– Hannover, 1716)[1] ranks with Newton as one of the inventors of the calculus. He was an infant prodigy, teaching himself Latin at the age of eight, and taking his degree in law before the age of twenty-one. In the service of the Elector of Mainz, and later in that of three successive dukes of Braunschweig-Lüneburg, he travelled extensively through England, France, Germany, Holland, Italy, everywhere seeking the acquaintance of prominent scholars. He finally settled at Hannover as librarian to the duke. In 1709 he was made a Baron of the Empire. When, in 1714, the Duke of Hannover crossed to England to become George I., he refused to allow Leibniz to accompany him. This embittered the last years of Leibniz's life.

His was a most versatile genius. He wrote on mathematics, natural science, history, politics, jurisprudence, economics, philosophy, theology, and philology. He invented a calculating machine that would add, subtract, multiply, divide, and even extract roots.

He was elected to membership in the Royal Society of London (1673), and to foreign membership in the Académie des Sciences (1700). He founded the Akademie der Wissenschaften (1700), and became its president for life. Many of his articles appear in the *Acta Eruditorum*, the organ of the last named society.[1]

His interest in the calculus must have been aroused while he was visiting England in 1672, where he probably heard from Oldenburg that Newton had some such method. His own development of the subject seems, however, to have been independent of that of Newton, while it shows the influence of both Barrow and Pascal.[2] He never published a work on the calculus, but confined himself to short articles in the *Acta Eruditorum*, and to piecemeal explanations of his discoveries in letters which he wrote to other mathematicians.

Clearly we are indebted to him for the following contributions to the development of the calculus:

1. He invented a convenient symbolism.
2. He enunciated definite rules of procedure which he called algorithms.
3. He realized and taught that quadratures constitute only a special case of integration; or, as he then called it, the inverse method of tangents.
4. He represented transcendental lines by means of differential equations.

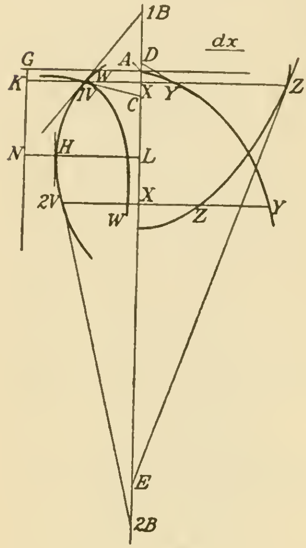
---

[1] For a brief sketch of the life of Leibniz see Smith, *History of Mathematics*, Vol. I., p. 417; also other histories of mathematics and the various encyclopedias.

[2] For example, his use of a characteristic triangle.

These points are illustrated in the following selections from two articles that were published in the *Acta Eruditorum*.[1]

The following extract is from "A new method for maxima and minima..." by Gottfried Wilhelm von Leibniz.[2]



Let there be an axis AX and several curves, as VV, WW, YY, ZZ, whose ordinates VX, WX, YX, ZX, normal to the axis, are called respectively, v, w, y, z; and the AX, cut off from the axis, is called x. The tangents are VB, WC, YD, ZE, meeting the axis in the points B, C, D, E, respectively. Now some straight line chosen arbitrarily is called dx, and the straight [line] which is to dx as v (or w, or y, or z) is to VB (or WC, or YD, or ZE), is called dv (or dw, or dy, or dz) or the difference of the v's (or the w's, or the y's, or the z's). These things assumed, the rules of the calculus are as follows:

If a is a given constant,

$$da = 0,$$

and

$$d\overline{ax} = adx;$$

if

$$y = v,$$

(or [if] any ordinate whatsoever of the curve YY [is] equal to any corresponding ordinate of the curve VV),

$$dy = dv.$$

Now, addition and subtraction: if

$$z - y + w + x = v,$$

$$d\overline{z - y + w + x} = dv,$$

or

$$= dz - dy + dw + dx.$$

---

[1] The Latin is frequently bad. The translator wishes to acknowledge her indebtedness to Professors Carter and Hahn, both of Hunter College, who kindly made a number of corrections.

[2] "Nova methodus pro maximis & minimis, itemque tangentibus, qua nec irrationales quantitates moratur, & singulare pro illis calculi genus, per G.G.L." (his Latin initials) *Acta Eruditorum*, October, 1684.

Multiplication:

$$d\overline{vx} = xdv + vdx,$$

or by placing

$$y = xv,$$
$$dy = xdv + vdx.$$

. . . .

Yet it must be noticed that the converse is not always given by a differential equation, except with a certain caution, of which [I shall speak] elsewhere.

Next, division:

$$d\frac{v}{y} = \frac{\pm\, vdy \mp ydv}{yy}$$

$\left(\text{or } z \text{ being placed equal to } \frac{v}{y}\right)$

$$dz = \frac{\pm\, vdy \mp ydv}{yy}$$

Until this sign may be correctly written, whenever in the calculus its differential is simply substituted for the letter, the same sign is of course to be used, and $+dx$ [is] to be written for $+x$, and $-dx$ [is] to written for $-x$, as is apparent from the addition and subtraction done just above; but when an exact value is sought, or when the relation of the z to x is considered, then [it is necessary] to show whether the value of the dz is a positive quantity, or less than nothing, or as I should say, negative; as will happen later, when the tangent ZE is drawn from the point Z, not toward A, but in the opposite direction or below X, that is, when the ordinates z decrease with the increasing abscissas x. And because the ordinates v sometimes increase, sometimes decrease, dv will be sometimes a positive, sometimes a negative quantity; and in the former case the tangent $1V1B$ is drawn toward A, in the latter $2V2B$ is drawn in the opposite direction. Yet neither happens in the intermediate [position] at M, at which moment the v's neither increase nor decrease, but are at rest; and therefore dv becomes equal to 0, where nothing represents a quantity [which] may be either positive or negative, for $+0$ equals -0; and at that place the v, obviously the ordinate LM, is maximum (or if the convexity turns toward the axis, minimum) and the tangent to the curve at M is drawn neither above X, where it approaches the axis in the direction of A, nor below X in the contrary direction, but is parallel to the axis. If dv is infinite

with respect to the dx, then the tangent is perpendicular to the axis, or it is the ordinate itself. If dv and dx [are] equal, the tangent makes half a right angle with the axis. If, with increasing ordinates v, their increments or differences dv also increase (or if, the dv's being positive, the ddv's, the differences of the differences are also positive, or [the dv's being] negative, [the ddv's are also] negative), the curve turns [its] convexity toward the axis; otherwise [its] concavity.[1] Where indeed the increment is maximum or minimum, or where the increments from decreasing become increasing, or the contrary, there is a point of opposite flexion, and the concavity and convexity are interchanged, provided that the ordinates too do not become decreasing from increasing or the contrary, for then the concavity or convexity would remain; but it is impossible that the amounts of change[2] should continue to increase or decrease while the ordinates become decreasing from increasing or the contrary. And so a point of flexion occurs when, neither v nor dv being 0, yet ddv is 0. Whence, furthermore, problems of opposite flexion have, not two equal roots, like problems of maximum, but three.

Powers:[3]

. . . .

$$dx^a = ax^{a-1}dx,$$

for example,

$$dx^3 = 3x^2dx;$$

$$d\frac{1}{x^a} = -\frac{adx}{x^{a-1}},$$

for example, if

$$w = \frac{1}{x^3},$$

$$dw = -\frac{3dx}{x^4}.$$

Roots:

$$d\sqrt[b]{x^a} = \frac{a}{b}\sqrt[b]{x^{a-b}}dx.$$

---

[1] [In the original the words concavity and convexity are interchanged, but farther on in the article the statement is made correctly.]

[2] [The word that Leibniz uses is "crementa."]

[3] [There are several mistakes in this paragraph and the following one but, as they are obviously printer's errors, they have been corrected by the translator.]

(Hence

$$d\sqrt[2]{y} = \frac{dy}{2\sqrt[2]{y}},$$

for in this case a is 1, and b is 2; therefore $\frac{a}{b}\sqrt[b]{x^{a-b}}$ is $\frac{1}{2}\sqrt[2]{y^{-1}}$; now y^{-1} is the same as $\frac{1}{y}$, from the nature of the exponents of a geometric progression, and $\frac{1}{\sqrt[2]{y}}$ is $\sqrt[2]{y^{-1}}$.)

$$d\frac{1}{\sqrt[b]{x^a}} = \frac{-adx}{b\sqrt[b]{x^{a+b}}}.$$

Again the rule for an integral power would suffice for determining fractions as well as roots, for the power may be a fraction while the exponent is negative, and it is changed into a root when the exponent is a fraction; but I have preferred to deduce these consequences myself rather than to leave them to be deduced by others, since they are completely general and of frequent occurrence; and in a matter which is itself involved it is preferable to take ease[1] [of operation] into account.

From this rule, known as an algorithm, so to speak, of this calculus, which I call differential, all other differential equations may be found by means of a general calculus, and maxima and minima, as well as tangents [may be] obtained, so that there may be no need of removing fractions, nor irrationals, nor other aggregates, which nevertheless formerly had to be done in accordance with the methods published up to the present. The demonstration of all [these things] will be easy for one versed in these matters, who also takes into consideration this one point which has not received sufficient attention heretofore, that dx, dy, dv, dw, dz, can be treated as proportional to the momentaneous differences, whether increments or decrements, of x, y, v, w, z (each in its order).[2]

. . . .

[1] [The word used is "facilitati."]

[2] [Leibniz now proceeds to illustrate his rules by means of a number of examples. These are followed by selections from the article "On abstruse geometry...," "De geometria recondita et analysi indivisibilium atque infinitorum, addenda bis qua sunt in *Actis* a. 1684, Maji, p. 233; Octob. p. 467; Decem. p. 585." G. G. L. *Acta Eruditorum*, June, 1686.

The errors in the pages to which reference is made have been corrected by the translator.]

Since, furthermore the method of investigating indefinite quadratures or their impossibilities is with me only a special case (and indeed an easier one) of the far greater problem which I call the *inverse method of tangents*, in which is included the greatest part of all transcendental geometry; and because it could always be solved algebraically, all things were looked upon as discovered;[1] and nevertheless up to the present time I see no satisfactory result from it; therefore I shall show how it can be solved no less than the indefinite quadrature itself. Therefore, inasmuch as algebraists formerly assumed letters or general numbers for the quantities sought, in such transcendental problems I have assumed general or indefinite equations for the lines sought, for example, the abscissa and the ordinate [being represented] by the usual x and y, my equation for the line sought is,

$$0 = a + bx + cy + exy + fx^2 + gy^2, \text{ etc.};$$

by the use of this indefinitely stated equation, I seek the tangent to a really definite line (for it can always be determined, as far as need be),[2] and comparing what I find with the given property of the tangents, I obtain the value[s] of the assumed letters, a, b, c, etc., and even establish the equation of the line sought, wherein occasionally certain [things] still remain arbitrary; in which case innumerable lines may be found satisfying the question, which was so involved that many, considering the problem as not sufficiently defined at last, believed it impossible. The same things are also established by means of series. But, according to the calculation to be effected, I use many things, concerning which [I shall speak] elsewhere. And if the comparison does not succeed, I decide that the line sought is not algebraic but transcendental.

Which being done, in order that I may discover the species of the transcendence (for some transcendentals depend upon the general section of a ratio, or upon logarithms, others upon the general section of an angle, or upon the arcs of a circle, others upon other more complex indefinite questions); therefore, besides the letters x and y, I assume still a third, as v, which signifies a transcendental quantity, and from these three I form the general

[1] [The Latin is, . . . & quod algebraice semper posset solvi, omnia reperta haberentur, & vero nihil adhuc de eo extare video satisfaciens, ...]

[2] [The Latin of this parenthesis is: (semper enim determinari potest, quousque assurgi opus sit). This is not good Latin, but the meaning is probably as we have given it.]

equation for the line sought, from which I look for the tangent to the line according to my method of tangents published in the *Acta*, October, 1684, which does not preclude[1] transcendentals. Thence, comparing what I discover with the given property of the tangents to the curve, I find, not only the assumptions, the letters a, b, c, etc., but also the special nature of the transcendental.

. . . .

Let the ordinate be x, the abscissa y, let the interval between the perpendicular and the ordinate. . .be p; it is manifest at once by my method that

$$pdy = xdx,$$

. . . .

which differential equation being turned into a summation becomes

$$\int pdy = \int xdx.$$

But from what I have set forth in the method of tangents, it is manifest that

$$d\overline{\tfrac{1}{2}xx} = xdx;$$

therefore, conversely,

$$\frac{1}{2}xx = \int xdx$$

(for as powers and roots in common calculation, so with us sums and differences or $\int$ and d, are reciprocals). Therefore we have

$$\int pdx = \frac{1}{2}xx. \qquad \text{Q. E. D.}$$

Now I prefer to employ dx and similar [symbols], rather than letters for them, because the dx is a certain modification of the x, and so by the aid of this it turns out that, since the work must be done through the letter x alone, the calculus obviously proceeds with its own[2] powers and differentials, and the transcendental relations are expressed between x and another [quantity]. For which reason, likewise, it is permissible to express transcendental

---

[1] [The Latin word is "moratur" which means, literally, "it does not linger," or "it does not take into consideration." But as the method given really does include the case of transcendentals, accuracy of translation must be sacrificed in the interest of truth.]

[2] [That is the powers and differentials of the x.]

lines by an equation; for example, if the arc is a, the versine x, then we shall have

$$a = \int dx : \sqrt{2x - xx},$$

and if y is the ordinate of the cycloid, then

$$y = \sqrt{2x - xx} + \int dx : \sqrt{2x - xx},$$

which equation perfectly expresses the relation between the ordinate y and the abscissa x. and from it all the properties of the cycloid can be demonstrated; and the analytic calculus is extended in this way to those lines which hitherto have been excluded for no greater cause than that they were believed unsuited[1] to it; also the Wallisian interpolations and innumerable other things are derived from this source.

. . . .

It befell me, up to the present a tyro in these matters, that, from a single aspect of a certain demonstration concerning the magnitude of a spherical surface, a great light suddenly appeared. For I saw that in general a figure formed by perpendiculars to a curve, and the lines applied ordinatewise to the axis (in the circle, the radii), is proportional to the surface of that solid which is generated by the rotation of the figure about the axis. Wonderfully delighted by which theorem, since I did not know that such a thing was known to others, I straightway devised the triangle which in all curves I call the characteristic [triangle], the sides of which would be indivisible (or, to speak more accurately, infinitely small) or differential quantities; whence immediately, with no trouble, I established countless theorems, some of which I afterward observed in the works of Gregory and Barrow.

. . . .

Finally I discovered the supplement of algebra for transcendental quantities, of course, my calculus of indefinitely small quantities, which, the differential as well as [that] of either summations or quadratures, I call, and aptly enough if I am not mistaken, the *analysis of indivisibles and infinites*, which having been once revealed, whatever of this kind I had formerly wondered about seems only child's play and a jest.

---

[1] [The Latin is "incapaces," literally, "incapable of."]

BERKELEY

Analyst and Its Effect upon the Calculus

(Selected and Edited by Professor Florian Cajori, University of California, Berkeley, Calif.)

By the publication of the *Analyst* in 1734, Dean (afterward Bishop) Berkeley profoundly influenced mathematical thought in England for more than half a century. His brilliant defence of his views against the criticisms of James Jurin of Cambridge and John Walton of Dublin, and the controversies among the mathematicians themselves which were started by Berkeley's *Analyst* (London, 1734), led to a clarifying of mathematical ideas as found in two important English books:—Benjamin Robin's *Discourse Concerning the Nature and Certainty of Sir Isaac Newtons Method of Fluxions and of Prime and Ultimate Ratios*, 1735, and Colin Maclaurin's *Treatise of Fluxions*, 1742. Robins greatly improved the theory of limits. Both Robins and Maclaurin banished from their works the fixed infinitesimal. In some recent books, one meets with the statement that it was Weierstrass who first banished the fixed infinitesimal from the calculus. This claim needs to be qualified by the historical fact that already in the eighteenth century the fixed infinitesimal was excluded from the works on the calculus written by Robins, Maclaurin, and also by Simon Lhuilier on the Continent.

The treatment of the calculus as initiated by Leibniz, became known in Great Britain earlier than the theory of fluxions. The Scotsman, John Craig, used the Leibnizian notation, in print, in 1685. The Newtonian fluxional notation was first printed in John Wallis's *Algebra* of 1693. Harris, Hayes, and Stone, though using the term "fluxion" and the notation of Newton, nevertheless drew their inspiration, on matters relating to mathematical concepts, from continental writers who followed Leibniz. These facts explain the reason why Berkeley devoted considerable attention to the calculus of Leibniz.

The *Analyst* is a book of 104 pages. It is addressed to "an infidel mathematician," generally supposed to have referred to the astronomer, Edmund Halley. There is no evidence of religious skepticism in Halley's published writings; his alleged "infidelity" rests only upon common repute. In the selections here given no effort has been made to preserve the capitalization of many of the words as in the original edition.

The Analyst:

A Discourse Addressed to an Infidel Mathematician

Though I am a stranger to your person, yet I am not, Sir, a stranger to the reputation you have acquired in that branch of

learning which hath been your peculiar study; nor to the authority that you therefore assume in things foreign to your profession; nor to the abuse that you, and too many more of the like character, are known to make of such undue authority, to the misleading of unwary persons in matters of the highest concernment, and whereof your mathematical knowledge can by no means qualify you to be a competent judge...

Whereas then it is supposed that you apprehend more distinctly, consider more closely, infer more justly, and conclude more accurately than other men, and that you are therefore less religious because more judicious, I shall claim the privilege of a Free-thinker; and take the liberty to inquire into the object, principles, and method of demonstration admitted by the mathematicians of the present age, with the same freedom that you presume to treat the principles and mysteries of Religion; to the end that all men may see what right you have to lead, or what encouragement others have to follow you...

The Method of Fluxions is the general key by help whereof the modern mathematicians unlock the secrets of Geometry, and consequently of Nature. And, as it is that which hath enabled them so remarkably to outgo the ancients in discovering theorems and solving problems, the exercise and application thereof is become the main if not the sole employment of all those who in this age pass for profound geometers. But whether this method be clear or obscure, consistent or repugnant, demonstrative or precarious, as I shall inquire with the utmost impartiality, so I submit my inquiry to your own judgment, and that of every candid reader.—Lines are supposed to be generated[1] by the motion of points, planes by the motion of lines, and solids by the motion of planes. And whereas quantities generated in equal times are greater or lesser according to the greater or lesser velocity wherewith they increase and are generated, a method hath been found to determine quantities from the velocities of their generating motions. And such velocities are called fluxions: and the quantities generated are called flowing quantities. These fluxions are said to be nearly as the increments of the flowing quantities, generated in the least equal particles of time; and to be accurately in the first proportion of the nascent, or in the last of the evanescent increments. Sometimes, instead of velocities, the momentaneous

[1] *Introd. ad Quadraturam Curvarum.*

increments or decrements of undetermined flowing quantities are considered, under the appellation of moments.

By moments we are not to understand finite particles. These are said not to be moments, but quantities generated from moments, which last are only the nascent principles of finite quantities. It is said that the minutest errors are not to be neglected in mathematics: that the fluxions are celerities, not proportional to the finite increments, though ever so small; but only to the moments or nascent increments, whereof the proportion alone, and not the magnitude, is considered. And of the aforesaid fluxions there be other fluxions, which fluxions of fluxions are called second fluxions. And the fluxions of these second fluxions are called third fluxions: and so on, fourth, fifth, sixth, etc., *ad infinitum.* Now, as our Sense is strained and puzzled with the perception of objects extremely minute, even so the Imagination, which faculty derives from sense, is very much strained and puzzled to frame clear ideas of the least particles of time, or the least increments generated therein: and much more so to comprehend the moments, or those increments of the flowing quantities in *statu nascenti,* in their very first origin or beginning to exist, before they become finite particles. And it seems still more difficult to conceive the abstracted velocities of such nascent imperfect entities. But the velocities of the velocities—the second, third, fourth, and fifth velocities, etc.—exceed, if I mistake not, all human understanding. The further the mind analyseth and pursueth these fugitive ideas the more it is lost and bewildered; the objects, at first fleeting and minute, soon vanishing out of sight. Certainly, in any sense, a second or third fluxion seems an obscure Mystery. The incipient celerity of an incipient celerity, the nascent augment of a nascent augment, *i. e.*, of a thing which hath no magnitude—take it in what light you please, the clear conception of it will, if I mistake not, be found impossible; whether it be so or no I appeal to the trial of every thinking reader. And if a second fluxion be inconceivable, what are we to think of third, fourth, fifth fluxions, and so on without end?...

All these points, I say, are supposed and believed by certain rigorous exactors of evidence in religion, men who pretend to believe no further than they can see. That men who have been conversant only about clear points should with difficulty admit obscure ones might not seem altogether unaccountable. But he who can digest a second or third fluxion, a second or third differ-

ence, need not, methinks, be squeamish about any point in divinity . . .

Nothing is easier than to devise expressions or notations for fluxions and infinitesimals of the first, second, third, fourth, and subsequent orders, proceeding in the same regular form without end or limit $\dot{x}$. $\ddot{x}$. $\dddot{x}$. $\ddddot{x}$ etc. or dx. ddx. $dddx$. $ddddx$. etc. These expressions, indeed, are clear and distinct, and the mind finds no difficulty in conceiving them to be continued beyond any assignable bounds. But if we remove the veil and look underneath, if, laying aside the expressions, we set ourselves attentively to consider the things themselves which are supposed to be expressed or marked thereby, we shall discover much emptiness, darkness, and confusion; nay, if I mistake not, direct impossibilities and contradictions. Whether this be the case or no, every thinking reader is entreated to examine and judge for himself. . .

This is given for demonstration.[1] Suppose the product or rectangle AB increased by continual motion: and that the momentaneous increments of the sides A and B are a and b. When the sides A and B were deficient, or lesser by one half of their moments, the rectangle was $\overline{A - \frac{1}{2}a} \times \overline{B - \frac{1}{2}b}$, *i. e.*, $AB - \frac{1}{2}aB - \frac{1}{2}bA + \frac{1}{4}ab$. And as soon as the sides A and B are increased by the other two halves of their moments, the rectangle becomes $\overline{A + \frac{1}{2}a} \times \overline{B + \frac{1}{2}b}$ or $AB + \frac{1}{2}aB + \frac{1}{2}bA + \frac{1}{4}ab$. From the latter rectangle subduct the former, and the remaining difference will be $aB + bA$. Therefore the increment of the rectangle generated by the entire increments a and b is $aB + bA$. Q. E. D. But it is plain that the direct and true method to obtain the moment or increment of the rectangle AB, is to take the sides as increased by their whole increments, and so multiply them together, $A + a$ by $B + b$, the product whereof $AB + aB + bA + ab$ is the augmented rectangle; whence, if we subduct AB the remainder $aB + bA + ab$ will be the true increment of the rectangle, exceeding that which was obtained by the former illegitimate and indirect method by the quantity ab. And this holds universally by the quantities a and b be what they will, big or little, finite or infinitesimal, increments, moments, or velocities. Nor will it avail to say that

[1] *Philosophiae Naturalis Principia Mathematica*, Lib. II, lem. 2.

ab is a quantity exceedingly small: since we are told that *in rebus mathematicis errores quam minimi non sunt contemnendi*[1]...

But, as there seems to have been some inward scruple or consciousness of defect in the foregoing demonstration, and as this finding the fluxion of a given power is a point of primary importance, it hath therefore been judged proper to demonstrate the same in a different manner, independent of the foregoing demonstration. But whether this method be more legitimate and conclusive than the former, I proceed now to examine; and in order thereto shall premise the following lemma:—"If, with a view to demonstrate any proposition, a certain point is supposed, by virtue of which certain other points are attained; and such supposed point be itself afterwards destroyed or rejected by a contrary

---

[1] *Introd. ad Quadraturam Curvarum.*

[Of interest are the remarks on Newton's reasoning, made in 1862 by Sir William Rowan Hamilton in a letter to Augustus De Morgan: "It is very difficult to understand the *logic* by which Newton proposes to prove, that the *momentum* (as he calls it) of the *rectangle* (or product) AB is equal to $aB + bA$, if the *momenta* of the sides (or factors) A and B be denoted by a and b. His mode of getting rid of ab appeared to me long ago (I must confess it) to involve so much of *artifice*, as to deserve to be called *sophistical;* although I should not like to say so publicly. He subtracts, you know $\left(A - \frac{1}{2}a)(B - \frac{1}{2}b\right)$ from $\left(A + \frac{1}{2}a\right)\left(B + \frac{1}{2}b\right)$; whereby, of course, ab disappears in the result. But by *what right*, or *what reason* other than to give an unreal air of *simplicity* to the calculation, does he *prepare* the *products* thus? Might it not be argued similarly that the difference,

$$\left(A + \frac{1}{2}a\right)^3 - \left(A - \frac{1}{2}a\right)^3 = 3aA^2 + \frac{1}{4}a^3$$

was the moment of A^3; and is it not a sufficient *indication* that the mode of procedure adopted is not the fit one for the subject, that it quite *masks* the notion of a *limit;* or rather has the appearance of treating that notion as foreign and irrelevant, notwithstanding all that had been said so well before, in the First Section of the First Book? Newton does not seem to have cared for being very consistent in his *philosophy*, if he could anyway get hold of *truth*-or what he considered to be such..." From *Life of Sir William Rowan Hamilton* by R. P. Graves, Vol. 3, p. 569.

We give also Weissenborn's objection to Newton's procedure of taking half of the increments a and b; with equal justice one might take, says he,

$$\left(A + \frac{2}{3}a\right)\left(B + \frac{2}{3}b\right) - \left(A - \frac{1}{3}a\right)\left(B - \frac{1}{3}b\right),$$

and the result would then be $Ab + Ba + \frac{1}{3}ab$. From H. Weissenborn's *Principien der höheren Analysis in ihrer Entwickelung von Leibniz bis auf Lagrange*, Halle, 1856, p. 42.]

supposition; in that case, all the other points attained thereby, and consequent thereupon, must also be destroyed and rejected, so as from thenceforward to be no more supposed or applied in the demonstration."[1] This is so plain as to need no proof.

Now, the other method of obtaining a rule to find the fluxion of any power is as follows. Let the quantity x flow uniformly, and be it proposed to find the fluxion of x^n. In the same time that x by flowing becomes $x + o$, the power x^n becomes $\overline{x + o|}^n$, *i. e.*, by the method of infinite series

$$x^n + nox^{n-1} + \frac{nn - n}{2} oox^{n-2} + \text{\&c.},$$

and the increments

$$o \text{ and } nox^{n-1} + \frac{nn - n}{2} oox^{n-2} + \text{\&c.}$$

are one to another as

$$1 \text{ to } nx^{n-1} + \frac{nn - n}{2} ox^{n-2} + \text{\&c.}$$

Let now the increments vanish, and their last proportion will be 1 to nx^{n-1}. But it should seem that this reasoning is not fair or conclusive. For when it is said, let the increments vanish, *i. e.*, let the increments be nothing, or let there be no increments, the former supposition that the increments were something, or that there were increments, is destroyed, and yet a consequence of that supposition, *i. e.*, an expression got by virtue thereof, is retained. Which, by the foregoing lemma, is a false way of reasoning. Certainly when we suppose the increments to vanish, we must

---

[1] [Berkeley's *lemma* was rejected as invalid by James Jurin and some other mathematical writers. The first mathematician to acknowledge openly the validity of Berkeley's *lemma* was Robert Woodhouse in his *Principles of Analytical Calculation*, Cambridge, 1803, p. XII. Instructive, in this connection, is a passage in A. N. Whitehead's *Introduction to Mathematics*, New York and London, 1911, p. 227. Whitehead does not mention Berkeley's *lemma* and probably did not have it in mind. Nevertheless, Whitehead advances an argument which is essentially the equivalent of Berkeley's, though expressed in different terms. When discussing the difference-quotient $\frac{(x + h)^2 - x^2}{h}$, Whitehead says: "In reading over the Newtonian method of statement, it is tempting to seek simplicity by saying that $2x + h$ is $2x$, when h is zero. *But this will not do; for it thereby abolishes the interval from x to $x + h$*, over which the average increase was calculated. The problem is, how to keep an interval of length h over which to calculate the average increase, and at the same time to treat h as if it were zero. Newton did this by the conception of a limit, and we now proceed to give Weierstrass's explanation of its real meaning."]

suppose their proportions, their expressions, and everything else derived from the supposition of their existence, to vanish with them...

I have no controversy about your conclusions, but only about your logic and method: how you demonstrate? what objects you are conversant with, and whether you conceive them clearly? what principles you proceed upon; how sound they may be; and how you apply them?...

Now, I observe, in the first place, that the conclusion comes out right, not because the rejected square of dy was infinitely small, but because this error was compensated by another contrary and equal error[1]...

The great author of the method of fluxions felt this difficulty, and therefore he gave in to those nice abstractions and geometrical metaphysics without which he saw nothing could be done on the received principles: and what in the way of demonstration he hath done with them the reader will judge. It must, indeed, be acknowledged that he used fluxions, like the scaffold of a building, as things to be laid aside or got rid of as soon as finite lines were found proportional to them. But then these finite exponents are found by the help of fluxions. Whatever therefore is got by such exponents and proportions is to be ascribed to fluxions: which must therefore be previously understood. And what are these fluxions? The velocities of evanescent increments. And what are these same evanescent increments? They are neither finite quantities, nor quantities infinitely small, nor yet nothing. May we not call them the ghosts of departed quantities...?

You may possibly hope to evade the force of all that hath been said, and to screen false principles and inconsistent reasonings, by a general pretence that these objections and remarks are *metaphysical.* But this is a vain pretence. For the plain sense and truth of what is advanced in the foregoing remarks, I appeal to the understanding of every unprejudiced intelligent reader...

And, to the end that you may more clearly comprehend the force and design of the foregoing remarks, and pursue them still farther in your own meditations, I shall subjoin the following Queries:—

Query 1. Whether the object of geometry be not the proportions of assignable extensions? And whether there be any need

[1] [Berkeley explains that the calculus of Leibniz leads from false principles to correct results by a "Compensation of errors." The same explanation was advanced again later by Maclaurin, Lagrange, and, independently, by L. N. M. Carnot in his *Réflexions sur la métaphysique du calcul infinitésimal,* 1797.]

of considering quantities either infinitely great or infinitely small? . . .

Qu. 4. Whether men may properly be said to proceed in a scientific method, without clearly conceiving the object they are conversant about, the end proposed, and the method by which it is pursued?. . .

Qu. 8. Whether the notions of absolute time, absolute place, and absolute motion be not most abstractely metaphysical? Whether it be possible for us to measure, compute, or know them? . . .

Qu. 16. Whether certain maxims do not pass current among analysts which are shocking to good sense? And whether the common assumption, that a finite quantity divided by nothing is infinite, be not of this number?[1]. . .

Qu. 31. Where there are no increments, whether there can be any *ratio* of increments? Whether nothings can be considered as proportional to real quantities? Or whether to talk of their proportions be not to talk nonsense? Also in what sense we are to understand the proportion of a surface to a line, of an area to an ordinate? And whether species or numbers, though properly expressing quantities which are not homogeneous, may yet be said to express their proportion to each other?. . .

Qu. 54. Whether the same things which are now done by infinites may not be done by finite quantities? And whether this would not be a great relief to the imaginations and understandings of mathematical men?. . .

Qu. 63. Whether such mathematicians as cry out against mysteries have ever examined their own principles?

Qu. 64. Whether mathematicians, who are so delicate in religious points, are strictly scrupulous in their own science? Whether they do not submit to authority, take things upon trust, and believe points inconceivable? Whether they have not *their* mysteries, and what is more, their repugnances and contradictions?". . .

---

[1] [The earliest exclusion of division by zero in ordinary elementary algebra, on the ground of its being a procedure that is inadmissible according to reasoning based on the fundamental assumptions of this algebra, was made in 1828, by Martin Ohm, in his *Versuch eines vollkommen consequenten Systems der Mathematik*, Vol. I, p. 112. In 1872, Robert Grassmann took the same position. But not until about 1881 was the necessity of excluding division by zero explained in elementary school books on algebra.

CAUCHY

On the Derivatives and Differentials of Functions of a Single Variable

(Translated from the French by Professor Evelyn Walker, Hunter College, New York City.)

Augustin-Louis Cauchy[1] (1789–1857), the well-known French mathematician and physicist, at the age of twenty-four gave up his chosen career as an engineer in order to devote himself to the study of pure mathematics. Soon afterward he became a teacher at the École Polytechnique. In 1816 he won the *Grand Prix* of the Institut for his mémoire on wave propagation. His greatest contributions to mathematics are embodied in the rigorous methods which he introduced. Of treatises and articles in scientific journals he published in all seven hundred and eighty-nine. His greatest achievement in the domain of the calculus was his scientifically correct derivation of the differential of a function, which he accomplished by means of the device that has come to be known as Cauchy's fraction. His treatment of the matter is as follows.[2]

Third Lesson

Derivatives[3] of Functions of a Single Variable

When the function $y = f(x)$ lies continuously between two given limits of x, and there is assigned to the variable a value included between these two limits, an infinitely small increment given to the variable produces an infinitely small increment of the function itself. Consequently, if we then place $\Delta x = i$, the two terms of the *ratio of the differences*

$$(1) \qquad \frac{\Delta y}{\Delta x} = \frac{f(x+i) - f(x)}{i}$$

will be infinitely small quantities. But while these terms indefinitely and simultaneously approach the limit zero, the ratio itself may converge toward another limit, either positive or nega-

[1] See C. A. Valson, *La Vie et les Traveaux du Baron Cauchy*, Paris, 1868.

[2] The two extracts quoted are from *Résumé des Leçons données à l'École Royale Polytechnique sur le Calcul Infinitésimal*, Paris, 1823. The present translation was made from the same work as republished in *Œuvres Complètes d'Augustin Cauchy*, Sér. II, Tome IV, Paris, 1889.

[3] [Cauchy's word is "derivées"]

tive. This limit, when it exists, has a fixed value for each particular value of x; but it varies with x. Thus, for example, if we take $f(x) = x^m$, m designating a whole number, the ratio between the infinitely small differences will be

$$\frac{(x+i)^m - x^m}{i} = mx^{m-1} + \frac{m(m-1)}{1.2}x^{m-2}i + \ldots + i^{m-1},$$

and it will have for [its] limit the quantity mx^{m-1}, that is to say, a new function of the variable x. It will be the same in general, only the form of the new function which serves as the limit of the ratio $\frac{f(x+i)-f(x)}{i}$ will depend upon the form of the given function $y = f(x)$. In order to indicate this dependence, we give to the new function the name *derived function*, and we designate it, with the help of an accent, by the notation

$$y' \text{ or } f'(x).$$

. . . .[1]

Fourth Lesson

Differentials of Functions of a Single Variable

Let $y = f(x)$ always be a function of the independent variable x; i an infinitely small quantity, and h a finite quantity. If we place $i = \alpha h$, α also will be an infinitely small quantity, and we shall have identically

$$\frac{f(x+i)-f(x)}{i} = \frac{f(x+\alpha h)-f(x)}{\alpha h},$$

whence there will result

$$(1) \qquad \frac{f(x+\alpha h)-f(x)}{\alpha} = \frac{f(x+i)-f(x)}{i}h.$$

The limit toward which the first member of equation (1) converges, while the variable α approaches zero indefinitely, the quantity h remaining constant, is what is called the *differential* of the function $y = f(x)$. We indicate this differential by the characteristic d, as follows:

$$dy \text{ or } df(x).$$

[1] [Cauchy then differentiates various functions using the above definition.]

It is easy to obtain its value when we know that of the derived function y' or $f'(x)$. Indeed, taking the limits of the two members of equation (1), we shall find generally

$$df(x) = hf'(x). \tag{2}$$

In the special case where $f(x) = x$, equation (2) reduces to

$$dx = h. \tag{3}$$

Therefore the differential of the independent variable x is nothing else than the finite constant h. That granted, equation (2) will become

$$df(x) = f'(x)dx \tag{4}$$

or, what amounts to the same thing,

$$dy = y'dx. \tag{5}$$

It follows from these last [equations] that the derived function $y' = f'(x)dx$ of any function $y = f(x)$ is precisely equal to dy/dx, that is to say, to the ratio between the differential of the function and that of the variable, or, if we wish, to the coefficient by which it is necessary to multiply the second differential in order to obtain the first. It is for this reason that we sometimes give to the derived function the name of *differential coefficient.*[1]

[1] [After this Cauchy gives the rules for differentiating various elementary functions, algebraic, exponential, trigonometric and antitrigonometric.]

EULER

On Differential Equations of the Second Order

(Translated from the Latin by Professor Florian Cajori, University of California, Berkeley, Calif.)

Euler's article from which we here quote represents the earliest attempt to introduce general methods in the treatment of differential equations of the second order. It was written when Leonhard (Léonard) Euler was in his twenty-first year and was residing in St. Petersburg, now Leningrad.

The title of the article is "A New Method of reducing innumerable differential equations of the second degree to differential equations of the first degree" (Nova methodvs innvmerabiles aeqvationes differentiales secvndi gradvs redvcendi ad aequationes differentiales primi gradvs). It was published in the *Commentarii academiae scientiarvm imperialis Petropolitanae*, Tomvs III ad annvm 1728, Petropoli, 1732, pp. 124–137.

When Euler in this article speaks of the "degree" of a differential equation, he means what we now call the "order" of such an equation. Observe also that he uses the letter c to designate 2.718..., the base of the natural system of logarithms. The first appearance (see page 95) of the letter e, in print, as the symbol for 2.178..., is in Euler's *Mechanica* (1733). We quote from Euler's article of 1728:

1. When analysts come upon differential equations of the second or any higher degree [order], they resort to two modes of solution. In the first mode they inquire whether it is easy to integrate them; if it is, they attain what they seek. When, however, an integration is either utterly impossible or at least more difficult, they endeavor to reduce them to differentials of the first degree, concerning which it is certainly easier to tell whether they can be resolved. Thus far no differential equations, save only those of the first degree, can be resolved by known [general] methods...

3. However, if in a differentio-differential equation one or the other of the indeterminates [variables] is absent, it is easy to reduce it to a simple differential, by substituting in the place of the differential of the missing quantity an expression composed of a new indeterminate multiplied by the other differential...As in the equation $Pdy^n = Qdv^n + dv^{n-2}ddv$, where P and Q signify any functions of y, and dy is taken to be constant. Since v does

not appear in the equation, let $dv = zdy$, then $ddv = dzdy$. Substituting these, yields the equation $Pdy^n = Qz^n dy^n + z^{n-2}dy^{n-1}dz$, and dividing by dy^{n-1}, gives the equation $Pdy = Qz^n dy + z^{n-2}dz$; this is a simple differential.

4. Except in this manner, no one, as far as I know, has thus far reduced other differentio-differential equations to differentials of the first degree, unless, perhaps, they admitted of being easily integrated directly. It is here that I advance a method by which to be sure not all, but numberless differentio-differential equations in whatever manner affected by any one variable, may be reduced to a simpler differential. Thus I am brought around to those reductions in which, by a certain substitution, I transform them [the differential equations] into others in which one of the indeterminates is wanting. This done, by the aid of the preceding section, the equations thus treated are reduced finally to differential equations of the first degree.

5. In this connection I observe this property of the exponential quantities, or rather powers, the exponent of which is variable, the quantity thus raised remaining constant, that if they are differentiated and differentiated again, the variable itself is restricted, so that it always affects only the exponent, and the differentials are composed of the integral itself multiplied by the differentials of the exponent. A quantity of this kind is c^x where c denotes the number, the logarithm of which is unity; its differential is $c^x dx$, its differentio-differential $c^x(ddx + dx^2)$, where x does not enter, except in the exponent. Considering these things, I observed that if in a differentio-differential equation, exponentials are thus substituted in place of the indeterminates, these variables remain only in the exponents. This being understood, these quantities must be so adapted, when substituted in place of the indeterminates that, after the substitution is made, they do not resist being removed by division; in this manner one in determinate or the other is eliminated and only its differentials remain.

6. This process is not applicable in all cases. But I have noticed that it holds for three types of differential equations of the second degree. The first type embraces all those equations which have only two terms...

7. All equations of the first type are embraced under this general formula: $ax^m dx^p = y^n dy^{p-2} ddy$, where dx is taken to be constant...To reduce that equation I place $x = c^{\alpha v}$, and $y = c^v t$. There result $dx = \alpha c^{\alpha v} dv$, and $dy = c^v(dt + tdv)$. And from this,

$ddx = \alpha c^{\alpha v}(ddv + \alpha dv^2)$ and $ddy = c^v(ddt + 2dtdv + tddv + tdv^2)$. But since dx is taken constant, one obtains $ddx = 0$, and $ddv = -\alpha dv^2$. Writing this in place of ddv, there follows $ddy = c^v. (ddt + 2dtdv + (1 - \alpha)tdv^2)$. Substituting these values in place of x and y in the given equation, it is transformed into the following, $ac^{\alpha v(m+p)}\alpha^p dv^p = c^{(n+p-1)v}t^n(dt + tdv)^{p-2}(ddt + 2dtdv + (1 - \alpha). tdv^2)$.

8. Now α should be so determined that the exponentials may be eliminated by division. To do this it is necessary that $\alpha v(m+p) = (n + p - 1)v$, whence one deduces $\alpha = \frac{n + p - 1}{m + p}\cdot$ Thus, α being determined, the above equation is changed to the following

$$a\left(\frac{n + p - 1}{m + p}\right)^p dv^p = t^n(dt + tdv)^{p-2}. \left(ddt + 2dtdv + \frac{m - n + 1}{m + p}tdv^2\right).$$

This may be deduced from the given equation directly, if I place $x = c^{(n+p-1)v:(m+p)}$, and $y = c^v t$. But $n + p - 1$ is the number of the dimensions which y determines; and $m + p$ which x determines. It is easy, therefore, in any special case to find α and to substitute the result. In the derived equation, since v is absent, place $dv = zdt$, then $ddv = zddt + dzdt$, but

$$ddv = -\alpha dv^2 = \frac{1 - n - p}{m + p}z^2dt^2.$$

From this follows $ddt = \frac{-dzdt}{z} + \frac{1 - n - p}{m + p}zdt^2$. After substituting these, there emerges

$$a\left(\frac{n + p - 1}{m + p}\right)^p z^p dt^p = t^n(dt + tzdt)^{p-2}\left(\frac{1 - n - p}{m + p}zdt^2 - \frac{dzdt}{z} + 2zdt^2 + \frac{m - n + 1}{m + p}tzzdt^2\right).$$

This divided by dt^{p-1} gives

$$a\left(\frac{n + p - 1}{m + p}\right)^p z^p dt = t^n(1 + tz)^{p-2}\left(\frac{1 + 2m - n + p}{m + p}zdt - \frac{dz}{z} + \frac{m - n + 1}{m + p}tz^2dt\right).$$

9. The proposed general equation $ax^m dx^p = y^n dy^{p-2} ddy$ is thus reduced to this differential of the first degree

$$a\left(\frac{n+p-1}{m+p}\right)^p z^{p+1}dt = t^n(1+tz)^{p-2}\left(\frac{1+2m-n+p}{m+p}z^2dt + \frac{m-n+1}{m+p}tz^3dt - dz\right),$$

the derived equation being multiplied by z. This equation may be obtained in one step from the one given, by placing in the first substitution $\int zdt$ in place of v. One should therefore take $x = c^{(n+p-1)\int zdt:(m+p)}$ and in place of y take $c^{\int zdt}t$; or what amounts to the same thing, place $x = c^{(n+p-1)\int zdt}$ and $y = c^{(m+p)\int zdt}t$...

10. We illustrate what we have derived in general terms by particular examples. Let $xdxdy = yddy$, which by division by dy, is reduced to $xdx = ydy^{-1}ddy$. Comparing this with the general equation one obtains $a = 1$, $m = 1$, $p = 1$, $n = 1$. Substituting these in the differential equation of the first degree, the given equation reduces to

$$\frac{1}{2}z^2dt = t(1+tz)^{-1}\left(\frac{3}{2}z^2dt + \frac{1}{2}tz^3dt - dz\right),$$

which becomes $z^2dt + tz^3dt = 3tz^2dt + t^2z^3dt - 2tdz$. The given equation, $xdxdy = yddy$, may be [directly] reduced to this by taking $x = c^{\int zdt}$ and $y = c^{2\int zdt}t$. Therefore, the construction [*i. e.*, resolution] of the proposed equation depends upon the construction of the derived differential equation...

11. The second type of differential equations which by my method I can reduce to differentials of the first degree, encompasses those which in the separate terms hold the same number of dimensions which the indeterminates and their differentials establish.[1] A general equation of this kind is the following: $ax^my^{-m-1}dx^pdy^{2-p} + bx^ny^{-n-1}dx^qdy^{2-q} = ddy$. In its separate terms the dimensions of the indeterminates [and their differentials] is unity. Also, dx is taken constant. This assumed equation is composed of only three terms, but as many as desired may be added to the above, the procedure remaining the same. There may be added $ex^ry^{-r-1}dx^qdy^{2-q}$ and as many of this kind as may be desired...

---

[1] [That is, the differential equations are homogeneous in x, y, dx, dy, and d^2y.]

12. I reduce the given equation by substituting c^v for x, and c^vt for y. Since therefore $x = c^v$ and $y = c^vt$, there follows $dx = c^vdv$ and $dy = c^v(dt + tdv)$; and from these, $ddx = c^v(ddv + dv^2)$ and $ddy = c^v(ddt + 2dtdv + tdv^2 + tddv)$. Since dx is taken to be constant, $ddx = 0$ and therefore $ddv = -dv^2$, and from this there results $ddy = c^v(ddt + 2dtdv)$. These values of x, y, dx, dy, and ddy, when placed in the equation, transform it into the following:

$$ac^vt^{-m-1}dv^p(dt + tdv)^{2-p} + bc^vt^{-n-1}dv^q(dt + tdv)^{2-q} = c^v(ddt + 2dtdv).$$

Dividing by c^v, this becomes

$$at^{-m-1}dv^p(dt + tdv)^{2-p} + bt^{-n-1}dv^q(dt + tdv)^{2-q} = ddt + 2dtdv.$$

Since v is absent from the equation, I place $dv = zdt$, and there will be $ddv = zddt + dzdt$, but $ddv = -dv^2 = -z^2dt^2$; therefore,

$$ddt = -zdt^2 - \frac{dzdt}{z}.$$

From this results the equation,

$$at^{-m-1}z^pdt^p(dt + ztdt)^{2-p} + bt^{-n-1}z^qdt^q(dt + ztdt)^{2-q} = - zdt^2 - \frac{dzdt}{z} + 2zdt^2,$$

or in better arrangement,

$$at^{-m-1}z^pdt(1 + zt)^{2-p} + bt^{-n-1}z^qdt(1 + zt)^{2-q} = zdt - \frac{dz}{z}.$$

13. This differential equation of the first degree may be derived from the given one by a single step, namely the straightway assumption that $x = c^{\int zdt}$ and $y = c^{\int zdt}t$. . .

18. The third type of equations, which I treat by this method of reducing, comprises those in which one or the other of the indeterminates in the separate terms hold the same number of dimensions. Here two cases are to be distinguished, according as the differential of the variable having everywhere the same dimension, is to be taken constant or not. To the first case[1] belongs the following general equation

$$Px^mdy^{m+2} + Qx^{m-b}dx^bdy^{m+2-b} = dx^mddy.$$

In this, x has the dimension m in each term, and dx is taken constant. Here P and Q signify any functions of y. For reducing this there

[1] [That is, the case in which the differential equation is homogeneous in x and dx.]

is need of only one substitution to wit, $x = c^v$, so that $dx = c^v dv$ and $ddx = c^v(ddv + dv^2) = 0$, and consequently, $ddv = -dv^2$. There results from this substitution

$$Pdy^{m+2} + Qdv^b dy^{m+2-b} = dv^m ddy,$$

of course, after dividing by c^{mv}.

19. Because of the absence of v in the derived equation, it can be reduced by substituting zdt for dv...

20. The other case of equations of the third type relates to the following general equation,

$$Px^m dy^{m+1} + Qx^{m-b} dx^b dy^{m-b+1} = dx^{m-1} ddx.$$

In this equation dy is taken constant, P and Q denoting any functions of y. And as one sees, x has the same dimension m in each term.[1] Take as before $x = c^v$; then $dx = c^v dv$, and $ddx = c^v(ddv + dv^2)$. When these are substituted in the equation, there results after division by c^{mv},

$$Pdy^{m+1} + Qdv^b dy^{m-b+1} = dv^{m+1} + dv^{m-1} ddv.$$

This equation is reduced as follows: Since v is absent, take $dv = zdy$, and, dy being constant, $ddv = dzdy$. Consequently, the last equation is changed to

$$Pdy^{m+1} + Qz^b dy^{m+1} = z^{m+1} dy^{m+1} + z^{m-1} dy^m dz.$$

But this, divided by dy^m, gives $Pdy + Qz^b dy = z^{m+1} dy + z^{m-1} dz$. Upon the reduction of this derived equation depends, therefore, the reduction of the given equation.

21. From this it will be understood, I trust, how differential equations of the second degree relating to one or another of the three types may be treated....

[1] [That is, the differential equation is homogenous in x, dx, and ddx.]

BERNOULLI

On the Brachistochrone Problem

(Translated from the Latin by Dr. Lincoln La Paz, National Research Fellow in Mathematics, The University of Chicago, Chicago, Ill.)

Jean (Johann, John) Bernoulli was born in Basel, Switzerland in 1667. He was professor of physics and mathematics at Groningen from 1695 until the chair of mathematics at Basel was vacated by the death of his elder brother, Jacques (Jakob, James) in 1705. Thereafter he was professor of mathematics at Basel until his own death in 1748. For further biographical details consult Merian, *Die Mathematiker Bernoulli* (Basel, 1860) or *Allgemeine Deutsche Biographie*, II, pp. 473–76.

The material translated in the following pages is collected in convenient form in Johannis Bernoulli, *Opera Omnia*, Lausanne and Geneva, 1742, vol. I, p. 161, pp. 166–169, pp. 187–193. The original sources are cited below in connection with the translations.

The calculus of variations is generally regarded as originating with the papers of Jean Bernoulli on the problem of the brachistochrone. It is true that Galileo in 1630–38[1] and Newton in 1686[2] had considered questions later recognized as belonging to the field of the calculus of variations. Their inquiries, however, are not looked upon as constituting the origin of this subject; since generality escaped them not only in the conception and formulation of their problems but also in the methods of attack which they devised.

On the contrary the writings of Jean Bernoulli show that he was not only fully aware of the difference between the ordinary problems of maxima and minima and the more difficult question he proposed, but also that he attained a fairly complete if not precise idea of the simpler problems of the calculus of variations *in general.* The terms in which he stated the problem of the brachistochrone may be readily extended to cover the formulation of the general case of the simplest class of variation problems in the plane. The curves he introduced under the name of *synchrones* for this problem furnish the first illustration of that important family of curves, now known as *transversals*, which is associated with the extremals of a problem in the calculus of variations; and in the fact noted by him, that the times of fall are equal along arcs intercepted by a synchrone on the cycloidal extremals of the brachistochrone problem which pass through a fixed point, we have the first instance of the beautiful *transversal theorem* of Kneser.[3]

[1] Galileo, *Dialog über die beiden hauptsächlichsten Weltsysteme* (1630) translation by Strauss, pp. 471–72; *Dialogues concerning Two New Sciences* (1638), translation by Crew and De Salvio, p. 239.

[2] Newton, *Principia*, Book II, Section VII, Scholium to Proposition XXXIV.

[3] Kneser, *Lehrbuch der Variationsrechnung*, 1900, p. 48;
Bolza, *Lectures on the Calculus of Variations*, University of Chicago Press, 1904, §33.

The reader of the translation which follows will note Jean Bernoulli's statement that he found a second or direct solution of the problem he proposed. In fact such a direct solution is mentioned in several of the letters which passed between Leibniz and Jean in 1696 as well as in the remarks which the former made on the subject of the brachistochrone problem in the *Acta Eruditorum* for May, 1697. However this direct demonstration which rests on the fundamental idea of general applicability employed by Jacques Bernoulli in obtaining his solution of the problem (namely that if a curve as a whole furnishes a minimum then the same property appertains to every portion of it) was not published until 1718 when both Jacques and Leibniz were dead. This fact is apparently regarded by those who believe Jean plagiarized from his brother Jacques as invalidating the former's claim of having secured a second solution. Jean for his part asserted that he delayed the publication of his second method in deference to counsel given by Leibniz in 1696.[1]

In any event it is regrettable that estimates of the relative value of the more mature methods of the two brothers often seem to be influenced by opinions which have been expressed with regard to the relative generality of their early solutions of the original brachistochrone problem, opinions which have in many cases been unfavorable to Jean Bernoulli. It is interesting to note in this connection that it was the opinion of as well qualified a student as Lagrange, if we may judge by statements made in his famous paper of 1762,[2] that all of the early solutions of the brachistochrone problem were found by special processes. In fact Lagrange emphasizes the part of Jean no less than that of Jacques in pioneering work on a general method in the calculus of variations.

New Problem

Which Mathematicians Are Invited to Solve[3]

If two points A *and* B *are given in a vertical plane, to assign to a mobile particle* M *the path* AMB *along which, descending under its own weight, it passes from the point* A *to the point* B *in the briefest time.*

To arouse in lovers of such things the desire to undertake the solution of this problem, it may be pointed out that the question proposed does not, as might appear, consist of mere speculation having therefore no use. On the contrary, as no one would readily believe, it has great usefulness in other branches of science such as mechanics. Meanwhile (to forestall hasty judgment) [it may be remarked that] although the straight line *AB* is indeed the shortest between the points *A* and *B*, it nevertheless is not the

---

[1] Consult in regard to this matter: Cantor, *Geschichte der Mathematik*, Vol. III, chap. 96, especially p. 226, p. 430, p. 439; Leibniz and Jean Bernoulli, *Commercium Philosophicum et Mathematicum*, Lausanne and Geneva, 1745, vol. I, p. 167, p. 178, especially p. 183 pp. 253–4, p. 266; Jean Bernoulli, *Opera Omnia*, vol. II, pp. 266–7.

[2] Lagrange, *Miscellanea Taurinensia*, vol. II, p. 173.

[3] [From the *Acta Eruditorum*, Leipzig, June, 1696, p. 269.]

path traversed in the shortest time. However the curve AMB, whose name I shall give if no one else has discovered it before the end of this year, is one well known to geometers.

. . . .

Proclamation

Made Public at Groningen, [Jan.], 1697

Jean Bernoulli public professor of mathematics pays his best respects to *the most acute mathematicians of the entire world.*

Since it is known with certainty that there is scarcely anything which more greatly excites noble and ingenious spirits to labors which lead to the increase of knowledge than to propose difficult and at the same time useful problems through the solution of which, as by no other means, they may attain to fame and build for themselves eternal monuments among posterity; so I should expect to deserve the thanks of the mathematical world if, imitating the example of such men as Mersenne, Pascal, Fermat, above all that recent anonymous Florentine enigmatist,[1] and others, who have done the same before me, I should bring before the leading analysts of this age some problem upon which as upon a touchstone they could test their methods, exert their powers, and, in case they brought anything to light, could communicate with us in order that everyone might publicly receive his deserved praise from us.

The fact is that half a year ago in the June number of the *Leipzig Acta* I proposed such a problem whose usefulness linked with beauty will be seen by all who successfully apply themselves to it. [An interval of] six months from the day of publication was granted to geometers, at the end of which, if no one had brought a solution to light, I promised to exhibit my own. This interval of time has passed and no trace of a solution has appeared. Only the celebrated Leibniz, who is so justly famed in the higher geometry has written[2] me that he has by good fortune solved this,

[1] Vincentius Viviani, A°. 1692. Aenigma Geometricum proposuit, d° miro opificio Testudinis quadrabilis Hemisphaericae; see *Acta Eruditorum* of this year, June, p. 274, or Vita Viviani in *Hist. Acad. Reg. Scient.*, Paris, A.e 1703. [The problems proposed by the other mathematicians referred to are well known.]

[2] [Leibniz and Jean Bernoulli, *Commercium Philosophicum et Mathematicum*, vol. I, p. 172. Leibniz in *Acta Erud.*, May, 1697, p. 202 credits Galileo with originally proposing the brachistochrone problem.]

as he himself expresses it, very beautiful and hitherto unheard of problem; and he has courteously asked me to extend the time limit to next Easter in order that in the interim the problem might be made public in France and Italy and that no one might have cause to complain of the shortness of time allotted. I have not only agreed to this commendable request but I have decided to announce myself the prolongation [of the time interval] and shall now see who attacks this excellent and difficult question and after so long a time finally masters it. For the benefit of those to whom the *Leipzig Acta* is not available, I here repeat the problem.

Mechanical—Geometrical Problem on the Curve of Quickest Descent.

To determine the curve joining two given points, at different distances from the horizontal and not on the same vertical line, along which a mobile particle acted upon by its own weight and starting its motion from the upper point, descends most rapidly to the lower point.

The meaning of the problem is this: Among the infinitely many curves which join the two given points or which can be drawn from one to the other, to choose the one such that, if the curve is replaced by a thin tube or groove, and a small sphere placed in it and released, then this [sphere] will pass from one point to the other in the shortest time.

In order to exclude all ambiguity let it be expressly understood that we here accept the hypothesis of Galileo, of whose truth, when friction is neglected, there is now no reasonable geometer who has doubt: *The velocities actually acquired by a heavy falling body are proportional to the square roots of the heights fallen through.* However our method of solution is entirely general and could be used under any other hypotheses whatever.

Since nothing obscure remains we earnestly request all the geometers of this age to prepare, to attack, to bring to bear everything which they hold concealed in the final hiding places of their methods. Let who can seize quickly the prize which we have promised to the solver. Admittedly this prize is neither of gold nor silver, for these appeal only to base and venal souls from which we may hope for nothing laudable, nothing useful for science. Rather, since virtue itself is its own most desirable reward and fame is a powerful incentive, we offer the prize, fitting for the man of noble blood, compounded of honor, praise, and approbation; thus we shall crown, honor, and extol, publicly and privately, in letter and by word of mouth the perspicacity of our great Apollo.

If, however, Easter passes and no one is discovered who has solved our problem, then we shall withhold our solution from the world no longer; then, so we hope, the incomparable Leibniz will permit to see the light his own solution and the one obtained by us which we confided to him long ago. If geometers will study these solutions which are drawn from deep lying sources, we have no doubt they will appreciate the narrow bounds of the ordinary geometry and will value our discovery so much the more, as so few have appeared to solve our extraordinary problem, even among those who boast that through special methods, which they commend so highly, they have not only penetrated the deepest secrets of geometry but also extended its boundaries in marvellous fashion; although their golden theorems, which they imagine known to no one, have been published by others long before.[1]

. . . .

The curvature of a beam of light in a non-uniform medium, and the solution of the problem proposed in the Acta 1696, p. 269, *of finding the* brachistochrone, *i. e., the curve along which a heavy particle slides down from given point to given point in the shortest time; and of the construction of the* synchrone, *or the wave-front of the beam.*[2]

Up to this time so many methods which deal with *maxima* and *minima* have appeared that there seems to remain nothing so subtle in connection with this subject that it cannot be penetrated by their discernment—so they think who pride themselves either as the originators of these methods or as their followers. Now the students may swear by the word of their master as much as they please, and still, if they will only make the effort, they will see that our problem cannot in any way be forced into the narrow confines imposed by their methods, which extend only so far as to determine a *maximum* or *minimum* among given quantities finite or infinite in number. Truly where the very quantities which are involved in our problem, from among which the maximum or minimum is to be found, are no more determinate than the very thing one is seeking—this is a task, this is difficult labor! Even those distinguished men, Descartes, Fermat, and others, who once contended as vigorously for the superiority of their

[1] [This remark is to be regarded as a covert thrust at Newton. As a matter of fact Newton, when the problem finally came to his attention, solved it immediately.]

[2] [From the *Acta Eruditorum*, Leipzig, May, 1697, p. 206.]

methods as if they fought for God and country[1] or in their place now their disciples, must frankly confess that the methods handed down from these same authorities are here entirely inadequate. It is neither my nature nor my purpose to ridicule the discoveries of others. These men certainly accomplished much and attained in admirable fashion the goal they had set for themselves. For just as in their writings we find no consideration whatever of this type of maxima and minima, so indeed they have not recommended their methods for any but common problems.

I do not propose to give a universal method, [a thing] that one might search for in vain; but instead particular methods of procedure by means of which I have happily unraveled this problem—methods which, indeed, are successful not only in this problem but also in many others. I decided to submit my solution immediately to the celebrated Leibniz, while others sought other solutions, in order that he might publish it together with his own in case he found one. That he would indeed find a solution I had no doubt, for I am sufficiently well acquainted with the genius of this most sagacious man. In fact, while I write this, I learn from one of the letters with which he frequently honors me that my problem had pleased him beyond [my] expectation, and (since it attracted him by its beauty, so he says, as the apple attracted Eve) he was immediately in possession of the solution. The future will show what others will have accomplished. In any case the problem deserves that geometers devote some time to its solution since such a man as Leibniz, so busy with many affairs, thought it not useless to devote his time to it. And it is reward enough for them that, if they solve it, they obtain access to hidden truths which they would otherwise hardly perceive.

With justice we admire Huygens because he first discovered that a heavy particle falls down along a *common cycloid* in the same time no matter from what point on the *cycloid* it begins its motion. But you will be petrified with astonishment when I say that precisely this *cycloid*, the *tautochrone of Huygens* is our required *brachistochrone*.[2] I arrived at this result along two different

[1] [For an interesting first hand account of this dispute between Fermat and Descartes on the subject of maxima and minima see the sequence of letters collected in *Œuvres de Fermat*, Paris, 1894, vol. II, pp. 126–168. See also page 610 of this Source Book.]

[2] [For a description of the cycloid and its properties see Teixeira, *Traité des Courbes Spéciales Remarquables*, 1909, vol. II, pp. 133–149, especially §540. See also R. C. Archibald, "Curves, Special, in the *Encyclopaedia Britannica*, 14th edition.]

paths, one indirect and one direct. When I followed the first [path] I discovered a wonderful accordance between the curved orbit of a ray of light in a continuously varying medium and our *brachistochrone curve.* I also observed other things in which I do not know what is concealed which will be of use in dioptrics. Consequently what I asserted when I proposed the problem is true, namely that it was *not mere speculation but would prove to be very useful in other branches of science*, as for example in dioptrics. But as what we say is confirmed by the thing itself, here is the first method of solution!

Fermat has shown in a letter to de la Chambre (see *Epist. Cartesii Lat.*, Tome III, p. 147, and Fermatii *Opera Mathem.* p. 156 et. seq.) that a ray of light which passes from a rare into a dense medium is bent toward the normal in such a manner that the ray (which by hypothesis proceeds successively from the source of light to the point illuminated) traverses the path which is shortest in time. From this principle he shows that the sine of the angle of incidence and the sine of the angle of refraction are directly proportional to the rarities of the media, or to the reciprocals of the densities; that is, in the same ratio as the velocities with which the ray traverses the media. Later the most acute Leibniz in *Act. Erud.*, 1682, p. 185 et. seq., and soon thereafter the celebrated Huygens in his treatise *de Lumine*, p. 40, proved in detail and justified by the most cogent arguments this same physical or rather metaphysical principle, which Fermat, contented with his geometric proof and all too ready to renounce the validity of his [least time] principle, seems to have abandoned under the pressure of Clerselier.[1]

[1] [The following remarks supply the historical background necessary for an appreciation of these statements of Bernoulli:

Fermat in a letter to de la Chambre in 1657 (see *Oeuvres de Fermat*, II, p. 354) emphasized his belief that Descartes had given no valid proof of his law of refraction (the law now credited to Snell). Fermat formulated his Least Time Principle in this letter and guaranteed that he could deduce from it all of the experimentally known properties of refraction by use of his method for solving problems of maxima and minima.

In 1662 Fermat, in compliance with a request made by de la Chambre, actually applied his Principle to the determination of the law of refraction. (*Oeuvres*, II, p. 457.) Since the Cartesian (Snell) law of refraction had been deduced by Descartes on the hypothesis that the velocity of light in a rare medium is less than in a dense medium, an assumption that Fermat regarded as obviously false; Fermat, employing the contrary hypothesis, looked forward with certainty to the discovery of a different law of refraction. To his amaze-

If we now consider a medium which is not uniformly dense but [is] as if separated by an infinite number of sheets lying horizontally one beneath another, whose interstices are filled with transparent material of rarity increasing or decreasing according to a certain law; then it is clear that a ray which may be considered as a tiny sphere travels not in a straight but instead in a certain curved path. (The above-mentioned Huygens notes this in his treatise *de Lumine* but did not determine the nature of the curve itself.) This path is such that a particle traversing it with velocity continuously increasing or diminishing in proportion to the rarity, passes from point to point in the shortest time. Since the sines of [the angles of] refraction in every point are respectively as the rarities of the media or the velocities of the particle it is evident also that the curve will have this property, that the sines of its [angles of] inclination to the vertical are everywhere proportional to the velocities. In view of this one sees without difficulty that the *brachistochrone* is the curve which would be traced by a ray of

---

ment he found on carrying through all the details of his minimizing process that the application of his Principle led to precisely the same law of refraction as that established by Descartes. Fermat was so confounded by this unexpected result that he agreed to cede the victory to the Cartesians; although his distrust of Descartes' mode of proof was manifest.

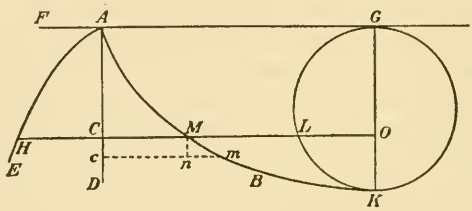
The Cartesian Clerselier impressed by the fact that if the Least Time Principle were true Descartes' hypothesis with regard to the velocity of light must be false (for he could find no error in Fermat's geometrical proof) applied himself zealously to overthrowing this principle. (*Œuvres*, II, letter CXIII, p. 464; letter CXIV, p. 472.)

Fermat (*Œuvres*, II, p. 483) apparently disgusted with the matter wrote in his answer to Clerselier:

"As to the principal question, it seems to me that I have often said not only to M. de la Chambre but to you that I do not pretend and I have never pretended to be in the secret confidence of nature. She moves by paths obscure and hidden which I have never made the attempt to penetrate. I have merely offered her a little aid from geometry in connection with the subject of refraction, in case this aid would be of use to her. But since you assure me that she can take care of her own affairs without this assistance, and that she is content to follow the path prescribed to her by M. Descartes, I abandon to you with all my heart my supposed conquest of physics [*i. e.* the Least Time Principle]; and I shall be content if you will leave me in possession of my problem of pure geometry *taken in the abstract*, by means of which we can find the path of a moving particle which passes through two different media and which seeks to achieve its motion in the shortest time." Fermat's renunciation of his Principle seems, however, to have been a transitory one; for, in 1664 we find him again attacking on the basis of this Principle Descartes's deduction of the law of refraction. (*Œuvres*, II, letter CXVI, p. 485.)]

light in its passage through a medium whose rarity is proportional to the velocity which a heavy particle attains in falling vertically. For whether the increase in the velocity depends on the nature of the medium, more or less resistant, as in the case of the ray of light, or whether one removes the medium, and supposes that the acceleration is produced by means of another agency but according to the same law, as in the case of gravity; since in both cases the curve is in the end supposed to be traversed in the shortest time, what hinders us from substituting the one in place of the other?

In this way we can solve our problem generally, whatever we assume to be the law of acceleration. For it is reduced to finding the curved path of a ray of light in a medium varying in rarity



arbitrarily. Let therefore FGD be the medium, bounded by the horizontal FG in which the radiating point A [is situated]. Let the vertical AD be the axis of the given curve AHE, whose associated HC determine the rarities of the medium at the heights AC, or the velocities of the ray, or corpuscle, at the points M. Let the curved ray itself which is sought be AMB. Call AC, x; CH, t; CM, y; the differential Cc, dx; diff. nm, dy; diff. Mm, dz; and let a be an arbitrary constant. Take Mm for the whole sine,[1] mn for the sine of the angle of refraction or of inclination of the curve to the vertical, and then by what we have just said, mn is to HC in constant ratio, that is $dy:t = dz:a$. This gives the equation $ady = tdz$, or $aady^2 = ttdz^2 = ttdx^2 + ttdy^2$; which when reduced gives the general differential equation $dy = tdx:\sqrt{(aa-tt)}$ for the required curve AMB. Thus I have with one stroke solved two remarkable problems, one optical the other mechanical, and [have accomplished] more than I required of others; I have shown that the two problems which are taken from entirely distinct fields of mathematics are nevertheless of the same nature.

Let us now consider a special case, namely that arising on the customary hypothesis first introduced and proved by Galileo,

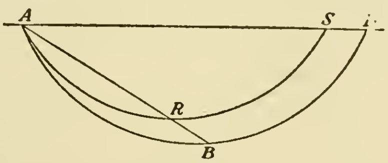
[1] [The Latin reading was changed from "*pro radio*" in the *Acta Erud.*, May 1697 to "*pro sinu toto*" in the *Opera Omnia*, 1742.]

according to which the velocity of heavy falling bodies varies as the square root of the distance fallen through; for this indeed is properly the problem. Under this assumption the given curve AHE will be a parabola, that is, $tt = ax$ and $t = \sqrt{ax}$. If this is substituted in the general equation we find $dy = dx\sqrt{\dfrac{x}{a - x}}$ from which I conclude that the *brachistochrone* curve is the *ordinary cycloid*. In fact if one rolls the circle GLK, whose diameter is a, on AG, and if the beginning of rotation is in A itself; then the point K describes a cycloid, which is found to have the same differential equation $dy = dx\sqrt{\dfrac{x}{a - x}}$, calling AC, x, and CM, y. Also this can be shown analytically from the preceding as follows:

$$dx\sqrt{\frac{x}{a - x}} = xdx : \sqrt{(ax - xx)}$$
$$= adx : 2\sqrt{(ax - xx)} - (adx - 2xdx) : 2\sqrt{(ax - xx)};$$

also $(adx - 2xdx) : 2\sqrt{(ax - xx)}$ is the differential quantity whose sum[1] is $\sqrt{(ax - xx)}$ or LO; and $adx : 2\sqrt{(ax - xx)}$ is the differential of the arc GL itself; and therefore, summing the equation $dy = dx\sqrt{\dfrac{x}{a - x}}$, we have y or $CM = GL - LO$, hence $MO = CO - GL + LO$. Since indeed (assuming CO = semiperiphery GLK) $CO - GL = LK$, we will have $MO = LK + LO$, and, cancelling LO, $ML = LK$; which shows the curve KMA to be the cycloid.

In order to completely solve the problem we have yet to show how from a given point, as vertex, we can draw the *brachistochrone*, or cycloid, which passes through a second given point. This is easily accomplished as follows: Join the two given points



A, B, by the straight line AB, and describe an arbitrary cycloid on the horizontal AL, having its initial point in A, and cutting the line AB in R; then the diameter of the circle which traces the required cycloid ABL passing through B is to the

[1] [Bernoulli uses sum for integral.]

diameter of the circle which traces the cycloid ARS as AB to AR.[1]

Before I conclude, I cannot refrain from again expressing the amazement which I experienced over the unexpected identity of *Huygens's tautochrone* and our *brachistochrone*. Furthermore I think it is noteworthy that this identity is found only under the hypothesis of Galileo so that even from this we may conjecture that nature wanted it to be thus. For, as nature is accustomed to proceed always in the simplest fashion, so here she accomplishes two different services through one and the same curve, while under every other hypothesis two curves would be necessary the one for oscillations of equal duration the other for quickest descent. If, for example, the velocity of a falling body varied not as the square root but as the cube root of the height [fallen through], then the *brachistochrone* would be algebraic, the *tautochrone* on the other hand transcendental; but if the velocity varied as the height [fallen through] then the curves would be algebraic, the one a circle, the other a straight line.[2]

---

[1] [Compare Bliss, *Calculus of Variations*, 1925, pp. 55–57.]

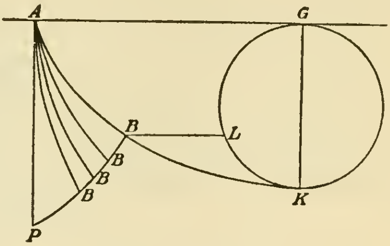
[2] [Denote by B_1 and B_2 the hypotheses which Bernoulli here suggests, and observe that he is concerned with a particle falling from rest in the velocity fields specified.

Under both B_1 and B_2, in case the initial velocity of the falling particle is different from zero, the integrand functions in the corresponding brachistochrone integrals, $T = \int_0^l \frac{ds}{v}$, are regular along the entire arc joining the given points A and B. On the contrary when the initial velocity is zero the integrand functions are singular at the initial point A and further investigation is necessary. For the corresponding case under the ordinary Galilean hypothesis consult Bliss, *Calculus of Variations*, 1925, p. 68; Kneser, *Lehrbuch der Variationsrechnung*, 2nd. Edt., 1925, p. 63.

The hypothesis B_2 is inadmissible in the case of the tautochrone. This may be shown by applying the method of Puiseux (see *Jour. de Math.*, [Liouville's], Ser. 1, vol. 9, p. 410; compare Appell, *Traité de Mécanique Rationelle*, vol. I, p. 351, and MacMillan, *Statics and the Dynamics of a Particle*, p. 225) to the integral $T = C\int_0^b \frac{\sqrt{1 + [f'(x)]^2}dx}{[b - x]^k}$ which represents the time of fall from the height $x = b$ to the height $x = 0$ along the curve whose equation is $y = f(x)$ when the velocity of fall is proportional to the kth power of the distance fallen through. We find in fact that for this integral to be independent of the value assigned to b it is necessary that $1 + [f'(x)]^2 = \delta^2 x^{2k-2}$. Consequently the integral is not well defined when $k = 1$.

This objection to the hypothesis B_2 was first raised and justified, essentially as above, by P. Stäckel in Oswald's *Klassiker der exakten Wissenschaften*, No. 46, 1894, Anmerkungen, p. 137. The formula for $f'(x)$ obtained in the

Geometers, I believe, will not be ungrateful if in conclusion I give the solution of a problem, just as worthy of consideration, which occurred to me while I was writing out what has gone before. *We require in the vertical plane the curve* PB, *which may be called the* Synchrone, *to every point* B *of which a heavy particle, descending from* A *along the cycloids* AB *with common vertex, arrives in the same time.* Let *AG* be horizontal and *AP* vertical. The meaning of the problem is as follows, that on each cycloid described on *AG*, a portion *AB* should be marked off, such that a heavy particle descending from *A* requires the same time to traverse it as it would require in falling from a given vertical height *AP*; when this is done, the point *B* will be in the *synchrone curve* PB which we seek.



If what we said above concerning light rays is considered attentively, it will be very evident that this curve is the same one which Huygens represents by the line *BC* in the figure on page 44 of his treatise *de Lumine*, and calls a *wave-front;* and just as the wave-front is cut orthogonally by all rays emanating from the light source *A*, as Huygens notes most opportunely; so our [curve] *PB* cuts all the cycloids with the point *A* as common vertex, at right angles. But if one had chosen to state the problem in this purely geometrical fashion: *to find the curve which cuts at right angles all the cycloids with common vertex;* then the problem would have been very difficult for geometers. However from the other point of view, regarding it as a falling body [problem], I construct [the curve] easily as follows. Let *GLK* be the generator circle of the cycloid *ABK*, and *GK* its diameter. Mark off the arc *GL* equal to the mean proportional between the given segment *AP* and the diameter *GK*. I say that *LB* drawn parallel to the horizontal *AG* cuts the cycloid *ABK* in the point *B*. If anyone wishes to try out his method on other [problems], let him seek the curve which cuts at right angles curves given successively in position (not indeed algebraic curves, for that would be by no means difficult, but transcendental curves), *e. g.*, logarithmic curves on a common axis and passing through the same point.

last paragraph shows that the equation of the tautochrone found by Stäckel is wrong. The error remains uncorrected in the revised (1914) edition of his article.]

ABEL

On Integral Equations

(Translated from the German by Professor J. D. Tamarkin, Brown University, Providence, R. I.)

The name of Niels Henrik Abel (b. August 5, 1802, d. April 6, 1829) deserves a place among those of the creators of our science, such as Newton, Euler, Gauss, Cauchy, and Riemann. During his short life Abel made numerous contributions to mathematics of the utmost importance and significance. Although his work was concentrated primarily on algebra and the integral calculus, his name will always be remembered in connection with many other branches of analysis, particularly the theory of integral equations, whose systematic development by Volterra, Fredholm and Hilbert began some 70 years after Abel's work.

We give here the translation of a short article under the title "Auflösung einer mechanischen Aufgabe," *Journal für die reine und angewandte Mathematik* (Crelle), Vol. I, 1826, pp. 153–157; *Œuvres Complètes*, Nouvelle édition par L. Sylow et S. Lie, Vol, I, 1881, pp. 97–101. This is a revised and improved version of an earlier paper: "Solution de quelques problèmes a l'aide d'intégrales définies" (in Norwegian), *Magazin for Naturvidenskaberne*, Aargang I, Bind 2, Christiania, 1823; *Œuvres Complètes.* Vol. I, pp. 11–18.

Abel solves here the famous problem of tautochrone curves by reducing it to an integral equation which now bears his name. His solution is very elegant and needs but slight modification to be presented in modern form. This solution and the formulas

$$\phi(x) = \int_0^\infty dq f(q) \cos qx, \qquad f(q) = \frac{2}{\pi}\int_0^\infty dx \phi(x) \cos qx$$

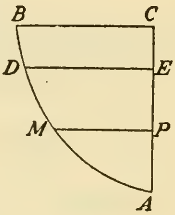
given by Fourier[1] are perhaps the first examples of an explicit

[1] "Théorie de mouvement de la chaleur dans les corps solides" *Mémoires de l'Académie royale des sciences de l'Institut de France*, Vol. 4, 1819–1820 (published in 1824) pp. 185–555 (489). This memoir was presented by Fourier in 1811 and was awarded a prize in 1812.

determination of an unknown function from an equation in which this function appears under the integral sign. An equation which can be reduced to that of Abel was given almost simultaneously by Poisson,[1] without solution. There exists now an extended literature devoted to Abel's equation and to analogous integral equations. The whole question is closely related to the notion of integrals and derivatives of non-integral order. The possibility of such operations was first suggested by Leibniz (1695) and Euler;[2] the notion was considerably developed by Liouville and Riemann,[3] and at present it has many important applications to various problems of pure and applied analysis.

Solution of a Mechanical Problem

Let $BDMA$ be any curve. Let BC be a horizontal and CA a vertical line. Let a particle move along the curve under the action of gravity, starting from the point D. Let τ be the time elapsed when the particle arrives at a given point A, and let a be the distance EA. The quantity τ is a function of a which depends upon the form of the curve and conversely, the form of the curve will depend upon this function. We shall investigate how is it possible to find the equation of the curve by means of a definite integral, if τ is a given continuous function of a.



Let $AM = s$, $AP = x$ and let t be the time in which the particle describes the arc DM. By the rules of mechanics we have

[1] "Second Mémoire sur la distribution de la chaleur dans les corps solides," *Journal de l'École Polytechnique*, Cahier 19, Vol. 12, 1823, pp. 249–403 (299).

[2] Leibniz, *Mathematische Schriften*, herausgegeben by C. I. Gerhardt, Halle Vol. 3, 1855 (letters to Johann Bernoulli), Vol. 4, 1859 (letters to Wallis).

L. Euler, "De progressionibus transcendentibus seu quarum termini generales algebraice dari nequeunt," *Commentarii Academiae Scientiarum Petropolitanae*, Vol. 5, 1730–1731, pp. 36–57 (55–57); *Opera Omnia* (1)14, pp. 1–25 (23–25).

[3] J. Liouville, "Mémoire sur quelques questions de géométrie et de mécanique, et sur un nouveau genre de calcul pour résoudre ces questions," *Journal de l'École Polytechnique*, Cahier 21, Vol. 13, 1832, pp. 1–69; "Mémoire sur le calcul des différentielles a l'indices quelconques," *ibidem*, pp. 71–162.

B. Riemann, "Versuch einer allgemeinen Auffassung der Integration und Differentiation," *Werke*, 2nd edition, 1892, pp. 353–366.

$-\frac{ds}{dt} = \sqrt{a-x}$,[1] whence $dt = -\frac{ds}{\sqrt{a-x}}$. Consequently, on integrating from $x = a$ to $x = 0$,

$$\tau = -\int_a^0 \frac{ds}{\sqrt{a-x}} = \int_0^a \frac{ds}{\sqrt{a-x}},$$

where $\int_\alpha^\beta$ means that the limits of integration correspond to $x = \alpha$ and $x = \beta$ respectively. Let now $\tau = \phi(a)$ be the given function; then

$$\phi(a) = \int_0^a \frac{ds}{\sqrt{a-x}}$$

will be the equation from which s is to be determined as a function of x. Instead of this equation we shall consider another, more general one,

$$\phi(a) = \int_0^a \frac{ds}{(a-x)^n}, \quad [0 < n < 1]$$

from which we shall try to derive the expression for s in terms of x.

If $\Gamma(\alpha)$ designates the function

$$\Gamma(\alpha) = \int_0^1 dx\left(\log\frac{1}{x}\right)^{\alpha-1}, \left[= \int_0^\infty e^{-z}z^{\alpha-1}dz\right]$$

t is known that

$$\int_0^1 y^{\alpha-1}(1-y)^{\beta-1}dy = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

where α and β must be greater than zero. On setting $\beta = 1 - n$ we find

$$\int_0^1 \frac{y^{\alpha-1}dy}{(1-y)^n} = \frac{\Gamma(\alpha)\Gamma(1-n)}{\Gamma(\alpha+1-n)},$$

whence, on putting $z = ay$,

$$\int_0^a \frac{z^{\alpha-1}dz}{(a-z)^n} = \frac{\Gamma(\alpha)\Gamma(1-n)}{\Gamma(\alpha+1-n)}a^{\alpha-n}.$$

---

[1] [If we designate by $v_0 = 0$ and v the velocities of the particle at the points D and M respectively and by g the acceleration due to gravity, then the equation of energy gives $v^2 - v_0^2 = 2g(a-x)$, whence

$$v = \frac{ds}{dt} = -\sqrt{2g}\sqrt{a-x}.$$

Thus the equation in the text corresponds to a choice of units such that $2g = 1$.]

Multiply by $da/(x - a)^{1-n}$ and integrate from $a = 0$ to $a = x$:

$$\int^{x} \frac{da}{(x-a)^{1-n}} \int_0^a \frac{z^{\alpha-1}dz}{(a-z)^n} = \frac{\Gamma(\alpha)\Gamma(1-n)}{\Gamma(\alpha+1-n)} \int_0^x \frac{a^{\alpha-n}da}{(x-a)^{1-n}}.$$

Setting $a = xy$ we have

$$\int_0^a \frac{a^{\alpha-n}da}{(x-a)^{1-n}} = x^\alpha \int_0^1 \frac{y^{\alpha-n}dy}{(1-y)^{1-n}} = x^\alpha \frac{\Gamma(\alpha-n+1)\Gamma(n)}{\Gamma(\alpha+1)};$$

hence

$$\int_0^x \frac{da}{(x-a)^{1-n}} \int_0^a \frac{z^{\alpha-1}dz}{(a-z)^n} = \frac{x^\alpha\Gamma(n)\Gamma(1-n)\Gamma(\alpha)}{\Gamma(\alpha+1)}$$

But, by a known property of the Γ-function,

$$\Gamma(\alpha+1) = \alpha\Gamma(\alpha),$$

whence, by substitution,

$$\int_0^x \frac{da}{(x-a)^{1-n}} \int_0^a \frac{z^{\alpha-1}dz}{(a-z)^n} = \frac{x^\alpha}{\alpha}\Gamma(n)\Gamma(1-n).$$

Multiplying this by $\alpha\phi(\alpha)d\alpha$ and integrating with respect to α [between any constant limits], we have

$$\int_o^x \frac{da}{(x-a)^{1-n}} \int_o^a \frac{(\int\phi(\alpha)\alpha z^{\alpha-1}d\alpha)dz}{(a-z)^n} = \Gamma(n)\Gamma(1-n)\int\phi(\alpha)x^\alpha d\alpha.$$

Setting

$$\int\phi(\alpha)x^\alpha d\alpha = f(x)$$

and differentiating, we have

$$\int\phi(\alpha)\alpha x^{\alpha-1}d\alpha = f'(x),\ \int\phi(\alpha)\alpha z^{\alpha-1}d\alpha = f'(z).$$

Then

$$\int_o^x \frac{da}{(x-a)^{1-n}} \int_o^a \frac{f'(z)dz}{(a-z)^n} = \Gamma(n)\Gamma(1-n)f(x)^1$$

---

1 [This identity follows immediately from the Dirichlet's formula

$$(\star)\qquad \int_0^x da\int_0^a F(a,z)dz = \int_0^x dz\int^x F(a,z)da \qquad (1)$$

(Bôcher, *An introduction to the study of integral equations*, 1909, p. 4) under certain restrictive assumptions as to $f(z)$. For instance, it suffices to assume that $f'(z)$ is continuous and $f(0) = 0$. Setting in $(\star)$

$$F(a,z) = (x-a)^{n-1}(a-z)^{-n}f'(z),$$

we see at once that the left-hand member of the equation in the text reduces to

$$\int_0^x f'(z)dz\int_z^x (x-a)^{n-1}(a-z)^{-n}da.$$

or, since

$$\Gamma(n)\Gamma(1-n) = \frac{\pi}{\sin n\pi},$$

$$(1) \qquad f(x) = \frac{\sin n\pi}{\pi}\int_0^x \frac{da}{(x-a)^{1-n}}\int_0^a \frac{f'(z)dz}{(a-z)^n}.$$

By means of this formula it is easy to find s from the equation

$$\phi(a) = \int_0^a \frac{ds}{(a-x)^n}.$$

Multiplying this equation by $\dfrac{\sin n\pi}{\pi}\dfrac{da}{(x-\alpha)^{1-n}}$ and integrating from $a = 0$ to $a = x$ we have

$$\frac{\sin n\pi}{\pi}\int_0^x \frac{\phi(a)da}{(x-a)^{1-n}} = \frac{\sin n\pi}{\pi}\int_0^x \frac{da}{(x-a)^{1-n}}\int^a \frac{ds}{(a-x)^n};$$

hence, by (1),

$$s = \frac{\sin n\pi}{\pi}\int_0^x \frac{\phi(a)da}{(x-a)^{1-n}}.$$ [1]

Substituting $a = z + t(x - z)$ in the interior integral, we reduce it to

$$\int_0^1 t^{-n}(1-t)^{n-1}dt = \Gamma(n)\Gamma(1-n)$$

with the final result

$$\int_0^x (x-\alpha)^{n-1}da\int_0^a(\alpha - z)^{-n}f'(z)dz = \Gamma(n)\Gamma(1-n)\int_0^x f'(z)dz = \Gamma(n)\Gamma(1-n)f(x).$$

Strictly speaking, the method used in the text establishes the identity in question only for the functions $f(x)$ which can be represented by definite integrals of the form

$$\int \phi(\alpha)x^\alpha d\alpha$$

but the investigation of the possibility of such a representation requires the solution of an integral equation of more complicated form than the given one.]

[1] [Two observations should be made concerning the solution obtained.

1. Since the function s replaces $f(x)$ of (1), it must satisfy the restrictions imposed upon $f(x)$, for instance $s'(x)$ must be continuous and $s(0) = 0$, which is natural in view of the physical interpretation of $s(x)$. This imposes certain restrictions upon the given function $\phi(a)$; it is easily verified that the conditions above are satisfied provided $\phi'(a)$ is continuous and $\phi(0) = 0$. The last condition again follows quite naturally from the integral equation of the problem.

2. If all the conditions above are satisfied, identity (1) shows immediately that the solution of the problem is unique, for, if $f(z)$ is a solution, then the interior integral in (1) reduces to $\phi(a)$, which yields the solution obtained in the text.]

Now let $n = \frac{1}{2}$; then

$$\phi(a) = \int_0^a \frac{ds}{\sqrt{a - x}}$$

and

$$s = \frac{1}{\pi}\int_0^x \frac{\phi(a)da}{\sqrt{x - a}}.$$

This equation gives s in terms of the abscissa x, and the curve therefore is completely determined.

We shall aply the expression above to some examples.

I. If

$$\phi(a) = \alpha_0 a^{\mu_0} + \alpha_1 a^{\mu_1} + \ldots + \alpha_m a^{\mu_m} = \Sigma\alpha a^{\mu}$$

the expression for s will be

$$s = \frac{1}{\pi}\int_0^x \frac{da}{\sqrt{x - a}}\sum \alpha a^{\mu} = \frac{1}{\pi}\sum\left(\alpha\int_0^x \frac{a^{\mu}da}{\sqrt{x - a}}\right).$$

Setting $a = xy$ we have

$$\int_0^x \frac{a^{\mu}da}{\sqrt{x - a}} = x^{\mu + \frac{1}{2}}\int_0^1 \frac{y^{\mu}dy}{\sqrt{1 - y}} = x^{\mu + \frac{1}{2}}\frac{\Gamma(\mu + 1)\Gamma(\frac{1}{2})}{\Gamma(\mu + \frac{3}{2})};$$

hence

$$s = \frac{\Gamma(\frac{1}{2})}{\pi}\sum \frac{\alpha\Gamma(\mu + 1)}{\Gamma(\mu + \frac{3}{2})}x^{\mu + \frac{1}{2}}$$

or, since $\Gamma(\frac{1}{2}) = \sqrt{\pi}$,

$$s = \sqrt{\frac{x}{\pi}}\left[\alpha_0\frac{\Gamma(\mu_0 + 1)}{\Gamma(\mu_0 + \frac{3}{2})}x^{\mu_0} + \ldots + \alpha_m\frac{\Gamma(\mu_m + 1)}{\Gamma(\mu_m + \frac{3}{2})}x^{\mu_m}\right].$$

If $m = 0$ and $\mu_0 = 0$, the curve in question is an isochrone, and we find

$$s = \sqrt{\frac{x}{\pi}}\,\alpha_0\,\frac{\Gamma(1)}{\Gamma(\frac{3}{2})} = \frac{2\alpha_0\sqrt{x}}{\pi}$$

which is known to be equation of a cycloid.[1]

---

[1] [We omit example II, where the function $\phi(a)$ is assumed to be given by different formulas in different intervals.

In the earlier article mentioned above, Abel gives the same final formula for the solution but bases his discussion on the assumption that s can be represented by a sum of terms of the form

$$s = \Sigma\alpha^{(m)}x^m.$$

He then discusses particular cases where the time of descent is proportional to a power of the vertical distance a or is constant (isochrone curve). At the

end of the article Abel gives a striking form to his solution by using the notation of derivatives and integrals of non-integral order. We define as the derivative of order α of a function $\psi(x)$ the expression

$$\frac{d^\alpha\psi(x)}{dx^\alpha} = D_x{}^\alpha\psi(x) = \begin{cases} \dfrac{1}{\Gamma(-\alpha)}\displaystyle\int_c^x (x-z)^{-\alpha-1}\psi(z)dz \text{ if } \alpha < 0, \\ \dfrac{d^p}{dx^p}D_x^{\alpha-p}\psi(x) \text{ if } p \text{ is an integer and } 0 \leqq p - 1 < \alpha \leqq p, \end{cases}$$

c being a constant which equals 0 in Abel's discussion. If we assume without proof that $D^\alpha D^\beta\psi = D^{\alpha+\beta}\psi$, then Abel's integral equation can be written as

$$\phi(x) = \Gamma(1-n)D_x{}^{n-1}D_x s = \Gamma(1-n)D_x^n s,$$

which can be solved immediately by the formula

$$s(x) = \frac{1}{\Gamma(1-n)}D_x^{-n}\phi = \frac{1}{\Gamma(1-n)\Gamma(n)}\int_0^x \phi(a)(x-a)^{n-1}da.$$

To justify this operation Abel proves the identity

$$D_x{}^{-n-1}D_x{}^{n+1}f = f,$$

which also can be derived from the identity (1) above.

In the particular case $n = \frac{1}{2}$ Abel writes the equation and its solution respectively as

$$\psi(x) = \sqrt{\pi}\,\frac{d^{\frac{1}{2}}s}{dx^{\frac{1}{2}}};\quad s = \frac{1}{\sqrt{\pi}}\frac{d^{-\frac{1}{2}}\psi}{dx^{-\frac{1}{2}}} = \frac{1}{\sqrt{\pi}}\int^{\frac{1}{2}}\psi(x)dx^{\frac{1}{2}}.$$

At the end of the article Abel remarks: "In the same fashion as I have found s from the equation

$$\psi(a) = \int_{x=0}^{x=a}\overline{(a-x)^n}$$

I have determined the function ϕ from the equation

$$\psi(a) = \int\phi(xa)f(x)dx$$

where ψ and f are given functions and the integral is taken between any limits [constant?]; but the solution of this problem is too long to be given here." This solution was never published by Abel.

It should be noted finally that Abel's equation and several others analogous to it were solved by Liouville by using the notion of derivatives and integrals of non-integral order (loc. cit.). Liouville's procedure is purely formal, and he seems to be unaware of Abel's results. It was also Liouville who solved the equation of Poisson mentioned above (Note sur la détermination d'une fonction arbitraire placée sous un signe d'intégration définie, Journal de l'École Polytechnique, Cahier 24, Vol. 15, 1835, pp. 55–60). Poisson's equation is

$$F(r) = \frac{1}{2}\sqrt{\pi}\,r^{n+1}\int_0^\pi \psi(r\cos\omega)\sin^{2n+1}\omega\,d\omega$$

where $F(r)$ is a given function and the unknown function $\psi(a)$ is assumed to be even, $\psi(-u) = \psi(u)$. Poisson's equation is reduced to Abel's type by using $(0, \pi/2)$ as the interval of integration and by making the substitution $\cos\omega = (z/x)^{\frac{1}{2}}$, $r^2 = x$.]

BESSEL

On His Functions

(Translated from the German by Professor H. Bateman, California Institute of Technology, Pasadena, Calif.)

Friedrich Wilhelm Bessel was born on July 22, 1784, and died on March 17, 1846. His father was Regierungssecretär and finally obtained the title of Justizrath. His mother was the daughter of pastor Schrader of Rehme. He married Johanna Hagen of a Königsberg family and had two sons and three daughters. At Olbers's desire and proposal Bessel took the position of inspector to the private Observatory of the Oberamtmann Schröter in Lilienthal. This was early in 1806 and from this date Bessel was an astronomer by profession and worked with great zeal. In his observational work he paid much attention to the planet Saturn. The portion of his work that seems to be best known is the experimental work on pendulums but his name has become famous on account of the work or the functions which now bear his name.

He was not the first to use these functions but he was certainly the first to give a systematic development of their properties and some tables for the functions of lowest order. Bessel considered only the functions of order i where i is an integer, but similar functions of non-integral order have been found to be of importance in applied mathematics. The literature of the subject is now quite vast and many differential equations have been solved in terms of Bessel functions. The values of numerous definite integrals can also be expressed in terms of these functions; indeed, the functions have been found to be so useful that the tables of Bessel have been greatly extended and books have been devoted entirely to the development of the properties. These books have given most mathematicians all the formulas they require and I believe that very few men turn to the original memoir. This, however, is still of much interest and well deserves a place among the most important contributions to the progress of mathematics. On account of its length the memoir is not given in full, the extracts consist of the preface and some portions relating to the properties of the functions.

The translation (pp. 667–669) is from his "Untersuchungen des Theils der planetarischen Störungen, welcher aus der Bewegung der Sonne entsteht" (Investigation of the portion of the planetary perturbations which arises from the motion of the sun) which appeared in the *Berlin Abbandlungen* (1824), and in his *Werke*, Bd. 1, pp. 84–109. The translation was checked by Morgan Ward, Research fellow in mathematics of California Institute of Technology, Pasadena.

The disturbance of the elliptic motion of one planet by another consists of two parts: one arises from the attraction of the disturbing planet on the disturbed planet; the other arises from the motion

of the sun which the disturbing planet produces. The two parts are combined in previous calculations of planetary disturbances but it is worth while to try to separate them. The latter can, in fact, as I shall show in the present memoir, be directly and completely evaluated and so deserves to be separated from the first for which the evaluation has so far not been made; the separation is indeed necessary if we wish to subject to a test the assumption, generally made so far, that the disturbing planet acts on the disturbed and the sun with equal mass.[1]

The two integrals $\int \cos i\mu.\cos \epsilon.d\epsilon$ and $\int \sin i\mu.\sin \epsilon.d\epsilon$ occurring in the first six formulae can easily be reduced to

$$\int \cos (h\epsilon - k \sin \epsilon)d\epsilon$$

where h denotes an integer; this last integer I shall denote by $2\pi I_k{}^h$. We have in fact

$$\int \cos i\mu.\cos \epsilon.d\epsilon = \int \cos i\mu[1 - (1 - e\cos \epsilon)]\frac{d\epsilon}{e}$$

$$= \frac{1}{e}\int \cos i\mu.d\epsilon - \frac{1}{e}\int \cos i\mu.d\mu$$

where the last part vanishes when taken between $\mu = 0$ and $\mu = 2\pi$ thus

$$\int \cos i\mu.\cos \epsilon.d\epsilon = \frac{2\pi}{e}I_{ie}^{i}$$

Furthermore

$$\int \sin i\mu.\sin \epsilon.d\epsilon = \int \cos i\mu.\cos \epsilon.d\epsilon - \int \cos (\epsilon + i\mu)d\epsilon.$$

or

$$\int \sin i\mu.\sin \epsilon.d\epsilon = \frac{2\pi}{e}I_{ie}^{i} - 2\pi I_{je}^{i+1}$$

The series expansion for $I_h{}^k$ is obtained in the way used in my memoir on Kepler's problem,[2] it is

$$I_k{}^h = \frac{(k/2)^h}{\pi(h)}\left\{1 - \frac{1}{h+1}\left(\frac{k}{2}\right)^2 + \frac{1}{1.2(h+1)(h+2)}\left(\frac{k}{2}\right)^4 - \dots\right\}$$

---

[1] [Bessel's fundamental equations are

$$\frac{r}{a}\cos \phi = \cos \epsilon - e \qquad (1)$$

$$\frac{r}{a}\sin \phi = \sqrt{1 - e^2}.\quad \sin \epsilon \qquad (2)$$

$$u = \epsilon - e\sin \epsilon, \qquad r = a(1 - e\cos \epsilon).]$$

[2] ["Analytische Auflösung der Kepler'sche Aufgabe." *Abhandlungen der Berliner Akademie der Wissenschaften, math. Cl.* (1816–17), p. 49; *Werke*, Bd. I, p. 17. The paper was read July 2, 1818. It was also communicated to Lindenau in a letter written in June, 1818. See *Zeitschr. für Astron.*, V., p. 367.]

where

$$\pi(h) = h!$$

Not only the equation for the center and the quantities $\cos\phi$, $\sin\phi$, $r\cos\phi$, $r\sin\phi$, $\frac{1}{r^2}\cos\phi$, $\frac{1}{r^2}\sin\phi$ lead on expansion to these definite integrals but this is also the case for

$$\log r,\ r^n,\ r^n\cos m\phi,\ r^n\sin m\phi,\ r^n\cos m\epsilon,\ r^n\sin m\epsilon,$$

whenever n and m are integers which may be positive, negative, or zero. Since most problems of physical astronomy lead to such expansions in series, a fuller knowledge of these integrals is desirable.

For brevity the four integrals, taken between 0 and 2π, will be denoted by symbols as follows:—

$$\frac{2\pi}{e}L = \int \cos i\mu.\cos\epsilon.d\epsilon;\ \frac{2\pi}{e}L' = \int \sin i\mu.\sin\epsilon.d\epsilon;$$

$$\frac{2\pi}{e}M = \int \frac{\cos i\mu.\cos\epsilon.d\epsilon}{1 - e\cos\epsilon};\ \frac{2\pi}{\epsilon}M' = \int \frac{\sin i\mu.\sin\epsilon.d\epsilon}{1 - e\cos\epsilon},$$

and we must first show that the expansions of the quantities mentioned involve these quantities.

We denote the coefficient of $\cos i\mu$ in the expansion of $\log r$ by H^i; the expansion being made so that the series runs over both positive and negative values of i. We have

$$2\pi H^i = \int \log r.\cos i\mu d\mu = \frac{1}{i}\log r.\sin i\mu - \frac{e}{i}\int \frac{\sin i\mu.\sin\epsilon.d\epsilon}{1 - e\cos\epsilon}$$

thus, with the exception of $i = 0$,

$$H^i = -\frac{1}{i}M'.$$

For $i = 0$ we obtain a logarithmic expansion; in fact, if we denote $\frac{e}{1+\sqrt{1-e^2}}$ by λ and take the semi-major axis to be unity we have

$$\frac{1}{r} = \frac{1}{\sqrt{1-e^2}}\left\{1 + 2\lambda\cos\epsilon + 2\lambda^2\cos 2\epsilon + 2\lambda^3\cos 3\epsilon + \ldots\right\};$$

and, if we multiply by $dr = e\sin\epsilon d\epsilon$ and integrate,

$$\log r = c - 2\{\lambda\cos\epsilon + \tfrac{1}{2}\lambda^2\cos 2\epsilon + \tfrac{1}{3}\lambda^3\cos 3\epsilon + \ldots\}.$$

For the determination of the constant c we have, for $\epsilon = 0$,

$$\log (1 - e) = c - 2\{\lambda + \tfrac{1}{2}\lambda^2 + \tfrac{1}{3}\lambda^3 + \ldots\}$$
$$= c + 2 \log (1 - \lambda).$$

$\therefore \log r =$

$$\log\frac{1 - e}{(1 - \lambda)^2} - 2\left\{\lambda \cos \epsilon + \frac{1}{2}\lambda^2 \cos 2\epsilon + \frac{1}{3}\lambda^3 \cos 3\epsilon + \ldots\right\};$$

and, if we multiply this by $d\mu = (1 - e \cos \epsilon)d\epsilon$ and integrate from 0 to 2π,

$$H^0 = \log \frac{1 - e}{(1 - \lambda)^2} + \lambda e = \log \frac{1 + \sqrt{1 - e^2}}{2} + \frac{e^2}{1 + \sqrt{1 - e^2}}$$

Bessel's Recurrence Formulas.

The following recurrence formulae are given in Bessel's paper and will be quoted here without proof

$$o = kI_k^{i-1} - 2iI_k^i + kI_k^{i+1} \tag{40}$$

$$I_k^{-i} = (-)^i I_k^i \tag{41}$$

$$\frac{I_k^i}{I_k^{i-1}} = \cfrac{\frac{k}{2i}}{1 - \cfrac{\frac{k^2}{(2i)(2i+2)}}{1 - \cfrac{\frac{k^2}{(2i+2)(2i+4)}}{\ldots\ldots\ldots\ldots}}}$$

$$\frac{dI_k^i}{dk} = \frac{i}{k}I_k^i - I_k^{i+1}$$

$$\frac{I_k^{i+h}}{\left(\frac{k}{2}\right)^{ith}} = (-)^h \frac{d^h\left\{\frac{I_k^i}{\left(\frac{k}{2}\right)^i}\right\}}{\left(d\frac{k^2}{4}\right)^h},$$

$$\frac{I_k^i}{\left(\frac{k}{2}\right)^i} = (-)^i \frac{d^i\{I_k^0\}}{\left\{d\frac{k^2}{4}\right\}^i}$$

$$o = \frac{d^2I_k^i}{dk^2} + \frac{1}{k}\frac{dI_k^i}{dk} + I_k^i\left(1 - \frac{i^2}{k^2}\right)$$

Other integrals can be reduced to the function $I_k{}^i$ as the following examples will show[1].

$$\frac{1}{2\pi}\int \cos\,(i\epsilon - m\cos\epsilon - n\sin\epsilon)\,d\epsilon = \cos i\alpha . I^i{}_{\sqrt{m^2+n^2}} \tag{50}$$

$$\frac{1}{2\pi}\int \cos i\epsilon . \cos\,(m\cos\epsilon + n\sin\epsilon)d\epsilon = \cos i\alpha . I^i{}_{\sqrt{m^2+n^2}} \quad i \text{ even} \tag{51}$$
$$= 0 \quad i \text{ odd}$$

$$\frac{1}{2\pi}\int \sin i\epsilon . \sin\,(m\cos\epsilon + n\sin\epsilon)d\epsilon = \cos i\alpha . I^i{}_{\sqrt{m^2+n^2}} \quad i \text{ odd} \tag{52}$$
$$= 0 \quad i \text{ even}$$

Bessel then proves the following relations

$$\frac{1}{2\pi}\int (\cos\epsilon)^{2i}\cos\,(k\sin\epsilon)d\epsilon = \frac{1.3\ldots(2i-1)}{k^i}I_k{}^i \tag{53}$$

$$\frac{1}{2\pi}\int_0^1 \cos kz.(1 - z^2)^{i-\frac{1}{2}}dz = \frac{1.3\ldots(2i-1)}{4k^i}I_k{}^i \tag{54}$$

$$\frac{1}{2\pi}\int e^{n\cos\epsilon}\cos\,(m\sin\epsilon)d\epsilon = I^o{}_{\sqrt{m^2-n^2}} \tag{55}$$

$$\cos k.I_k{}^o = 1 - \frac{3}{(\pi 2)^2}k^2 + \frac{3.5.7}{(\pi 4)^2}k^4 - \ldots$$

where $(\pi 2)$ is written for $2!$, and so for similar uses of π.

$$\sin k.I_k{}^o = k - \frac{3.5}{(\pi 3)^2}k^3 + \frac{3.5.7.9}{(\pi 5)^2}k^5 - \ldots \tag{56}$$

$$I_{k+z}{}^i = \left[\left(1 + \frac{z}{k}\right)^i \left\{ I_k{}^i - I_k{}^{i+1}z\left((1 + \frac{z}{2k}\right) + \frac{I_k{}^{i+2}}{1.2}z^2\left(1 + \frac{z}{2k}\right)^2 - \ldots\right] \tag{49}$$

The last series can be used for the calculation and interpolation of a table of the functions and was actually used by Bessel to form tables of $I_k{}^o$ and $I_k{}^1$ from $k = 0$ to $k = 3.2$.

The Roots of the Bessel Functions.

"The function $I_k{}^o$ has in common with the sine and cosine the remarkable property of vanishing twice and changing sign when its argument k increases from $2n\pi$ to $(2n + 2)\pi$. I shall show that if

m is even $I_k{}^o > 0$ from $k = m\pi$ to $(m + \frac{1}{2})\pi$
" " odd $I_k{}^o < 0$ " " " "

[1] In these equations $\tan\alpha = n/m$.

If we put $\sin \epsilon = z$ and $k = \frac{2m + m'}{2}\pi$ where m denotes a proper fraction, we have, according to the remark (54)

$$I_{k^o} = \frac{2}{\pi}\int_0^1 \cos\left(\frac{2m + m'}{2}\pi z\right)\frac{dz}{\sqrt{1 - z^2}}$$

Writing v for $(2m + m')z$ this expression becomes

$$I_{k^o} = \frac{2}{\pi}\int_0^{2m+m'} \cos\frac{\pi v}{2}\,\frac{dv}{\sqrt{(2m + m')^2 - v^2}}$$

The integral taken from $v = a$ to $v = b$, when we write $b + u$ for v, is

$$\int_{a-b}^{b-b} \cos\left(\frac{b\pi}{2} + \frac{\pi u}{2}\right)\frac{du}{\sqrt{(2m + m')^2 - (b + u)^2}}$$

Taking successively $b = 1, 3, \ldots 2m - 1$ and a, b always $b - 1$, $b + 1$ respectively, the last expression gives

$$I_k{}^0 = \frac{2}{\pi}\int_{-1}^1 \sin\frac{\pi u}{2}\,du\left[-\frac{1}{\sqrt{\mu^2 - (1 + u)^2}} + \frac{1}{\sqrt{\mu^2 - (3 + u)^2}}\right.$$
$$\left. - \ldots + \frac{(-)^{m-1}}{\sqrt{\mu^2 - (2m - 3 + u)^2}} + \frac{(-)^m}{\sqrt{\mu^2 - (2m - 1 + u)^2}}\right\}$$
$$+ \frac{2}{\pi}(-)^m \int_0^{m'} \frac{\cos\frac{\pi u}{2}du}{\sqrt{(\mu^2 - 2m + u)^2}} (\mu = 2m + m')$$

The individual terms of this expression are +, the last clearly so because $\pi u/2$ is always $<\pi/2$, the other because their part is greater than the negative; for we have

$$\int_{-1}^1 \frac{\sin\frac{\pi\mu}{2}du}{\sqrt{\mu^2-(b+u)^2}} = \int_0^1 \left\{ \frac{\sin\frac{\pi u}{2}du}{\sqrt{\mu^2-(b+u)^2}} - \frac{\sin\frac{\pi u}{2}du}{\sqrt{\mu^2-(b-u)^2}} \right\}$$

where the denominator of the positive part is always smaller than that of the negative. Furthermore, each following term is greater than the preceding on account of the continually decreasing denominator; the sum of two successive terms has therefore the sign of the last. If m is even, the last term in the bracket is positive and therefore the sum of all terms positive; if m is odd the last term is negative and therefore the sum of all terms up to the second negative and the first term as well as the term outside the bracket is negative. This property does not belong to I_{k^o} alone but all the

I_k^i possess a similar property. In fact from (46) if for brevity we write $I_k^i = \frac{k}{2}R^{(i)}$ and $\frac{k^2}{4} = k$

$$R^{(i+1)} = -\frac{dR^i}{dx}$$

Therefore R^{i+1} vanishes when R_i has a maximum or minimum; but between two values of k or x, for which R^i vanishes there is necessarily a maximum or minimum, thus also a vanishing R^{i+1}. It is therefore clear that I_k^1 is zero just as often as I_k^o is a maximum or minimum; between two values of k for which $I_k^1 = 0$ there lies always a maximum or minimum of R^1, therefore a root of I_k^2 and so on."

MÖBIUS

On the Barycentric Calculus

(Translated from the German by J. P. Kormes, Hunter College, New York City.)

August Ferdinand Möbius (1790–1868) was professor of astronomy in Leipzig and wrote several papers on this subject. His researches in celestial mechanics led him to an extensive study of geometry. In 1827 he published his most important contribution to science under the title: *Der barycentrische Calcul ein neues Huelfsmittel zur analytischen Behandlung der Geometrie dargestellt und insbesondere auf die Bildung neuer Classen von Aufgaben und die Entwickelung mehrerer Eigenschaften der Kegelschnitte angewendet,* Leipzig, Verlag von Johann Ambrosius Barth," pp. 1–454.

In this work Möbius introduces for the first time homogeneous coordinates into analytic geometry. With the aid of the barycentric calculus the treatment of various problems and in particular those relating to conic sections becomes simple and uniform. He introduces the remarkable classification of the properties of geometric figures according to the transformations (similar, affine, collinear) under which these properties remain invariant. Möbius arrives at the characteristic invariant of the collinear group, the anharmonic ratio. He also succeeds in establishing the most general principle of duality of points and straight lines without the use of a conic section.

§2. Through two given points A and B parallel lines are drawn. If a and b are any two numbers in a given ratio such that $a + b$ is different from zezo, find a straight line intersecting the two parallel lines in A' and B' respectively such that

$$a.AA' + b.BB' = 0.$$

Draw the line AB and find on it a point P such that $AP:PB = b:a$. Every line through P (and no other line) intersecting the two parallel lines will have the required property. From the similarity of the triangles $AA'P$ and $BB'P$ we have

$$AA':BB' = AP:BP = AP:-PB^{1} = b:-a;$$

[1] [Möbius considers directed segments and triangles. Thus if A and B are any two points on a straight line, $AB + BA = 0$; and if B, C, D are three points on a straight line and A is a point not on the line, then the sum of the areas of the triangles

$$ACD + ADB + ABC = 0.]$$

hence

$$aAA' + bBB' = 0.$$

. . . .

§3, *c*. If we place in A and B weights proportional to a and b respectively, P may be considered as the centroid of the points A and B with the coefficients a and b.

. . . .

§8...THEOREM.—Given a system of n points $A, B, C, \ldots N$ with the coefficients $a, b, c, \ldots n$ respectively where the sum $a + b + c + \ldots + n$ is different from zero, there can always be found one point and only one point, the centroid S, having the following property: If parallel lines be drawn through the given points and through the point S in any direction and these lines be intersected by any plane in the points $A', B', C', \ldots S'$ respectively, we always have

$$a.AA' + b.BB' + \ldots + n.NN' = (a + b + \ldots + n).SS'.$$

In particular if the plane passes through S we have

$$a.AA' + b.BB' + \ldots + n.NN' = 0.$$

§9...If $a + b + c + \ldots + n = 0$, the centroid is infinitely remote in the direction determined by the parallel lines.

. . . .

§13...In place of the segments $AA', BB', \ldots$ their endpoints $A, B, \ldots$ shall be used. Thus if S is the centroid of A, B, C with the coefficients $a, b, -c$, we write

$$a.A + b.B - cC = (a + b - c).S.$$

. . . .

§14. The operations with such abbreviated formulas form the barycentric calculus or a calculus based upon the notion of the centroid...§15. (1) In barycentric calculus points and their coefficients are considered. The points are denoted by capital letters, their coefficients by small letters...(2) The fact that S is the centroid of the points $A, B, C, \ldots$ with the coefficients $a, b, c, \ldots$ is expressed as follows:

$$\text{I. } aA + bB + cC + \ldots = (a + b + c + \ldots)S[1]\ldots$$

(3) The fact that the system $A, B, C, \ldots$ with the coefficients $a, b, c, \ldots$ has the same centroid as the system $F, G, H, \ldots$ with the

---

[1] [$aA + bB + cC + \ldots \equiv S$ is used in place of I.]

coefficients f, g, h,...provided the sum of the coefficients a, b, c,... equals the sum of the coefficients f, g, h,...is expressed as follows:

$$\text{II. } aA + bB + cC + \ldots = fF + gG + hH + \ldots$$

(4) The equation III. $aA + bB + \ldots = 0$ indicates that the system A, B,...with the coefficients a, b,...has no finite centroid ...[1]

§21. Theorem.—If $aA + bB \equiv C$ then C is on the line through A and B and we have:

$$a:b = BC:CA\ldots$$

§23. Theorem.—If $aA + bB + cC \equiv D$ and A, B, C are not on a straight line, the point D is in the plane A, B, C, and so

$$a:b:c = DBC:DCA:DAB^{2}\ldots$$

§24...If $aA + bB + cC + dD \equiv 0$ then A, B, C, D are in one plane and we have:

$$a:b:c:d = BCD:-CDA:DAB:-ABC\ldots$$

§25. Theorem.—If $aA + bB + cC + dD \equiv E$ and A, B, C, D are not in one plane then we have:

$$a:b:c:d = \text{pyramids } BCDE:CDEA:DEAB:EABC\ldots$$

§28. In order to determine the position of a point be it on a straight line, plane or space quantities of two kinds are essential; the ones of the first kind remain the same for all points like the axes of the usual system of coordinates, the others, the coordinates in the most general sense, depend upon the position of the various points with respect to the quantities of the first kind. By the method under consideration points shall be determined as follows: The quantities of the first kind shall be points and we shall call them "fundamental points" and the point whose position is to be determined shall be considered as their centroid. These fundamental points are taken as the system of coordinates. The coordinates of any point P with respect to these fundamental points are given by the relations which must exist among the coefficients of the fundamental points in order that the point P should be the centroid of these points.

....

[1] [All equations in the barycentric calculus assume one of the forms I, II or III and retain this form throughout all transformations.]

[2] [DBC means the area of $\triangle$ DBC, etc.]

§36. The change from one system of coordinates to another is very simple. If A', B',..., the new fundamental points, are given in terms of A, B,..., then it is sufficient to express the old fundamental points A, B,...in terms of A', B',...If these values be substituted in the expression for P, P is given in terms of the new coordinates. The simplicity of this process is illustrated by the following example: Let A', B', C', the new fundamental points, be the midpoints of the sides of the fundamental triangle ABC, A' the midpoint of BC,[1] B' the midpoint of CA, and C' the midpoint of AB. We have then

$$2A' = B + C, 2B' = C + A, 2C' = A + B,$$

and therefore

$$A = B' + C' - A', B = C' + A' - B', C = A' + B' - C'.$$

If the expression for P with respect to the system ABC is

$$P \equiv pA + qB + rC,$$

then the expression for P with respect to the new system $A'B'C'$ becomes

$$P \equiv p(B' + C' - A') + q(C' + A' - B') + r(A' + B' - C'),$$

or

$$P \equiv (q + r - p)A' + (r + p - q)B' + (p + q - r)C'.$$

§144...Given a system of points A, B, C,..., three of these may be taken as the fundamental points and any other point in the plane will be determined if the ratios of the coefficients $a:b:c$ are given. If in another system of points A', B', C'...the fundamental triangle formed by A', B', C' has the same sides as the triangle ABC and the ratios $a':b':c'$ are equal to $a:b:c$ for every point, then any figure formed by the points in the second system will be equal and similar to the figure formed by the corresponding points in the first system. If the sides of the fundamental triangle $A'B'C'$ are not equal but are proportional to the sides of the triangle ABC, the corresponding figures are similar. Now let us assume that the ratios of the coefficients of corresponding points are equal but the choice of the fundamental triangle arbitrary.

§145...In order to study the relationship between corresponding figures take any three points A', B', C' as the new fundamental points corresponding to the points A, B, C respectively. Should

---

[1] [See page 671. According to I. we have: $bB + bC = (b + b)S = 2bS$; put $S = A'$.]

any point D' correspond to the point $D \equiv aA + bB + cC$ the expression for D'must be:

$$D' \equiv aA' + bB' + cC'$$

that is the area of the triangles formed by the points A', B', C', D' must be in the same proportion as the areas of the triangles formed by the points A, B, C, D...It follows that $A'B'C' = m.ABC$. This holds true for any two corresponding triangles...Since every figure may be considered as an aggregate of triangles, the nature of the relations under consideration is revealed in the fact that the areas of any two parts in one figure are to each other as the areas of the two corresponding parts in the other figure...

§147...The two figures are then said to be affine...

§153...In general all relations and properties of a figure which are expressed in terms of the coefficients of the fundamental points remain the same in all affine figures. Thus if in one figure two lines are parallel or if they intersect in a given point, the corresponding lines in the affine figure will be parallel or will intersect in the corresponding point. On the other hand all relations which cannot be expressed in terms of the coefficients of the fundamental points are different in affine figures...

§217...Consider now a relationship under the sole condition that straight lines correspond to straight lines and planes to planes...This relationship may be characterized as follows: A correspondence is set up between the points of two planes such that if in one plane a set of points coincides with a straight line [collineantur] the corresponding points in the second plane lie on a straight line. Hence the name for this relation is "collineation" ...

§200...Connect any four points A, B, C, D in a plane by straight lines. The resulting three points of intersection A', B', C' (Fig. 1) connect again by straight lines thus obtaining six new points of intersection: A'', B'', C'', F, G and H which in turn may be again connected with each other and with the seven points previously obtained by straight lines etc. The system of lines thus obtained from any four points A, B, C, D shall be called a *plane net* and the points A, B, C, D shall be called the *four fundamental points of the net.*

§201...Theorem.—If A, B, C, and $D = aA + bB + cC$ are the four fundamental points of the plane net, every point P of the net can be represented as follows:

$$P \equiv \varphi aA + \chi bB + \psi cC$$

where φ, χ, ψ are rational numbers including zero, which depend only upon the construction of the point P and not upon the four fundamental points.

. . . .

§202. . . THEOREM.—Every anharmonic ratio formed in a plane net is rational and depends only upon the construction of the straight lines and not upon the four fundamental points.

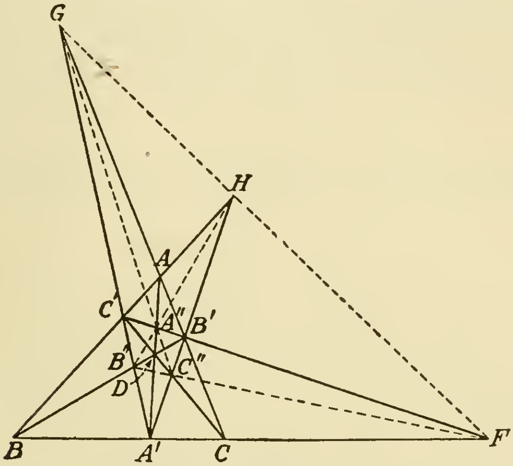
§219. Every point P of the net in the plane A, B, C, and $D = aA + bB + cC$ can be written in the form

$$P \equiv \varphi aA + \chi bB + \psi cC,$$

where φ, χ, ψ do not depend upon the coefficients a, b, c. Therefore every point P' of the net formed from the four fundamental points A', B', C', and $D' \equiv a'A' + b'B' + c'C'$ may be expressed as

$$P' \equiv \varphi a'A' + \chi b'B' + \psi c'C',$$

where φ, χ, ψ are the same as in the expression for the point P.



Since on the other hand every point of the plane ABC can be expressed in the form: $\varphi aA + \chi bB + \psi cC$ and when the values of φ, χ, and ψ are given it is always possible to find the point by mere drawing of straight lines, the relationship of collineation may now be defined as follows: Let any four points A', B', C', D', no three of which are on a straight line, correspond to four given points A, B, C, D, no three of which are on a straight line. To every point P in the first plane there will correspond a point P' in the second plane such that if

$$D \equiv aA + bB + cC, \qquad P \equiv pA + qB + rC,$$
$$D' \equiv a'A' + b'B' + c'C', P' \equiv p'A' + q'B' + r'C'.$$

Then,

$$\frac{p}{a}:\frac{q}{b}:\frac{r}{c} = \frac{p'}{a'}:\frac{q'}{b'}:\frac{r'}{c'}(= \varphi:\chi:\psi)\ldots$$

§220. If we determine the four pairs of the corresponding points A and A', B and B', C and C', D and D' then there corresponds one and only one point P' in the plane $A'B'C'$ to a point P in the plane ABC. The points A, B, C, D, P determine the ratios $a:b:c$ and $\varphi:\chi:\psi$, and the points A', B', C', D' determine the ratios $a':b':c'$. From these, the ratios $\varphi a':\chi b':\psi c'$ can be found and thus P' is uniquely determined. In the expression of two corresponding points

$$P \equiv \varphi aA + \chi bB + \psi cC, \text{ and } P' \equiv \varphi a'A' + \chi b'B' + \chi c'C',$$

if

$$\varphi a + \chi b + \psi c = 0, \varphi a' + \chi b' + \psi c'$$

may be different from zero. Thus a finite point may correspond to a point at infinity...

§221, 4. The collineation is characterized by the consistency of the ratios $\varphi:\chi:\psi$ for each pair of corresponding points. These ratios $\varphi:\chi:\psi$ can be expressed geometrically as anharmonic ratios. From

$$D \equiv aA + bB + cC \text{ and } P \equiv \varphi aA + \chi bB + \psi cC$$

it follows

$$a:b = -BCD:CDA$$

and

$$\varphi a:\chi b = -BCP:CPA$$

hence

$$\varphi:\chi = (B, A, CP, CD) = (A, B, CD, CP)$$

and similarly

$$\chi:\psi = (B, C, AD, AP)$$

(§190...where (A, B, CE, DE) is the anharmonic ratio the four points of which are A, B and the intersections of the line AB with the lines CE and DE respectively...). Therefore collineation may be defined directly by means of the equality of anharmonic ratios: Two figures are said to be collinear if every expression of the form:

$$\frac{ACD}{CDB}:\frac{AEF}{EFB}$$

is equal to the same expression formed by the corresponding points in the second figure.

SIR WILLIAM ROWAN HAMILTON

On Quaternions

(Selections Edited by Dr. Marguerite D. Darkow, Hunter College, New York City.)

William Rowan Hamilton was born in Dublin in 1805. His early training was in languages. At the age of thirteen, a copy of Newton's *Universal Arithmetic* fell into his hands and turned his thoughts to mathematics. In 1827, although an undergraduate, he was appointed to the chair of Astronomy in Trinity College, Dublin.

After producing a number of papers on various subjects, Hamilton concentrated upon the calculus of directed line-segments in space, and the meaning to be assigned to their product and their quotient. In 1835 and 1843, he wrote on this and allied matters in the *Transactions* of the Royal Irish Academy, and in 1844 in the *Philosophical Magazine*. In 1853, he published his *Lectures on Quaternions*. In 1866, his *Elements of Quaternions* (from which several extracts are to be quoted) appeared posthumously.

Although Hamilton expected his quaternions to prove a tool powerful for the progress of physics, his expectation has not been completely fulfilled, perhaps on account of a loss of naturalness in taking the square of a vector to be a negative scalar. The importance of his quaternions is due rather to the extension through them of the concept of number and the possibility of a variation from the hitherto unchallenged "Laws of Algebra."

Hamilton entered upon the mathematical scene at a time when mathematicians were not yet satisfied as to the sense in which $\sqrt{-1}$ was to be taken as a number, and had not quite divested it of all stigma for being "imaginary." Hamilton, too, was preoccupied with negative and imaginary, and rationalized these concepts by his view of algebra "as being *primarily* the science of ORDER (in time and space), and *not* primarily the science of MAGNITUDE." In his paper of 1835,[1] he introduces couples of moments in time (A, B), the difference A − B of two such moments as being a step a in time, couples of such steps (a, b), and couples of real numbers (a, b) which may be regarded as operators upon couples of steps. He defines

$$(a, b)(\mathrm{a}, \mathrm{b}) = (a\mathrm{a} - b\mathrm{b},\ a\mathrm{b} + b\mathrm{a}),$$
$$a(\mathrm{a}, \mathrm{b}) = (a, 0)(\mathrm{a}, \mathrm{b}) = (a\mathrm{a},\ a\mathrm{b}).$$

Hence $(0, 1)^2 (\mathrm{a}, \mathrm{b}) = (0, 1)(-\mathrm{b}, \mathrm{a}) = (-\mathrm{a}, -\mathrm{b}) = -(\mathrm{a}, \mathrm{b})$, whence $(0, 1)^2 = -1$.

From this beginning, there followed in the paper of 1848[2] a generalization to n-tuples (and in particular to quadruples) of moments and of steps in time.

---

[1] *Transactions* of the Royal Irish Academy, XVII, p. 293.

[2] [*Ibid XXI, p. 199.*]

He introduced the "momental quaternion" (A, B, C, D), A, B, C, D being moments of time,—the "ordinal quaternion" (a, b, c, d) = (a_0, a_1, a_2, a_3) with time-steps as elements, and the 16 $\times$ 24 operators $R_{\pm\pi,\pm\rho,\pm\sigma,\pm\tau}$ (π, ρ, σ, τ being some permutation of the integers, 0, 1, 2, 3) such that $R_{\pi\rho\sigma\tau}$ $(a_0, a_1, a_2, a_3) = (a_\pi, a_\rho, a_\sigma, a_\tau)$;

e. g. R_{3012} (a, b, c, d) = (d, a, b, c), $R_{-3,\ 0,\ 1,\ -2}$ (a, b, c, d) = ($-$d, a, b, $-$c)

He defined i as $R_{-1,\ 0\ -3,\ 2}$, j as $R_{-2,\ 3,\ 0,\ -1}$, k as $R_{-3,\ -2,\ 1,\ 0}$, whence followed the equations in operators:

$$i^2 = j^2 = k^2 = -1,\ ij = k = -ji,\ jk = i = -kj,\ ki = j = -ik,$$

with their non-commutative multiplication. Then came geometric interpretations and applications.

In the same paper, Hamilton laid a foundation for the modern work in linear algebras, regarded as a study of the properties of the set of linear combinations of n linearly independent elements with real or complex coefficients, subject to a multiplication table containing n^3 arbitrary constants, which closed the set under multiplication. Peacock (1791–1858) and DeMorgan (1806–1871) had already recognized the possibility of algebras which differ from ordinary algebra. Such algebras had appeared, but had found little recognition. It was the geometrical application of Hamilton's quaternions that led to a general appreciation of new algebras, in which the laws of combinations of the elements need not retain the classical commutativity, associativity, etc.

Hamilton was much concerned with ensuring respectability for his quaterions. The method of approach in his previous publications did not satisfy him. In his last work, he elected from the outset the geometric point of view. He defined a vector as a directed line-segment in space, and set himself the task of interpreting the quotient of two vectors, making the following assumptions (in which Greek letters denote vectors):

1. $\beta/\alpha = q$ implies $\beta = q\alpha$, in the sense that q, operating upon α, produces β.
2. $\beta'/\alpha' = \beta/\alpha$, $\alpha' = \alpha$, imply $\beta' = \beta$.
3. $q' = q$, $q'' = q$, imply $q' = q''$.
4. $\dfrac{\gamma}{\alpha} \pm \dfrac{\beta}{\alpha} = \dfrac{(\gamma \pm \beta)}{\alpha}$; $\dfrac{\gamma/\alpha}{\beta/\alpha} = \dfrac{\gamma}{\beta}$.
5. $\dfrac{\gamma}{\beta}\cdot\dfrac{\beta}{\alpha} = \dfrac{\gamma}{\alpha}$.

The quotient of two parallel vectors is plus or minus the ratio of their lengths, according to whether they are similarly or oppositely directed.

Three extracts from the *Elements of Quaternions*, London, 1866, are appended. The first (pp. 106–110) examines, on the basis of the preceding assumptions, the nature of the quotient of two vectors, and justifies the use of "quaternion" for such a quotient. The second (pp. 157–160) defines i, j, k and derives their multiplication table; and the third (pp. 149–150) comments on the source of the multiplicational non-commutativity.

"108. Already we may see grounds for the application of the *name* QUATERNION, to such a *Quotient of two Vectors* as has been spoken of in recent articles. In the first place, such a quo-

tient cannot *generally* be what we have called a SCALAR: or in other words, it cannot generally be equal to any of the (so-called) *reals of algebra*, whether of the *positive* or of the *negative* kind. For let x denote any such (actual)[1] scalar, and let α denote any (actual) vector; then we have seen that the product $x\alpha$ denotes *another* (actual) *vector*, say β', which is either *similar* or *opposite* in direction to α, according as the scalar coefficient, or *factor*, x, is positive or negative; in *neither* case, then, can it represent any vector, such as β, which is *inclined* to α, at any actual *angle*, whether acute, or right, or obtuse: or in other words, the *equation* $\beta' = \beta$, or $x\alpha = \beta$, is impossible, under the conditions here supposed. But we have agreed to write, as in algebra, $(x\alpha)/\alpha = x$; we must therefore[2]. . .*abstain* from writing *also* $\beta/\alpha = x$, under the same conditions: x still denoting a *scalar*. Whatever *else* a *quotient of two inclined vectors* may be found to be, it is thus, at least, a NON-SCALAR.

"109. Now, in forming the conception of the *scalar itself*, as the *quotient of two parallel*[3] *vectors*, we took into account not only *relative length*, or *ratio* of the usual kind, but also *relative direction*, under the form of *similarity or opposition*. In passing from α to $x\alpha$, we *altered* generally the *length* of the line α, in the ratio of $\pm x$ to 1; and we *preserved or reversed* the *direction* of that line, according as the *scalar coefficient* x was *positive* or *negative*, and, in like manner, in proceeding to form, more definitely than we have yet done, the conception of the *non-scalar quotient*, $q = \beta : \alpha = OB:OA$, of *two inclined vectors*, which for simplicity may be supposed to be *co-initial*, we have *still* to take account both of the *relative length* and of the *relative direction*, of the two lines compared. But while the *former* element of the *complex relation* here considered, between these two lines or vectors, is *still* represented by a simple RATIO (of the kind commonly considered in geometry), or by a *number*[4] expressing that ratio; the *latter element* of the same complex relation is *now* represented by an ANGLE, AOB: and not simply (as it was before) by an *algebraical sign* $+$ or $-$.

"110. Again, in estimating this *angle*, for the purpose of *distinguishing* one quotient of vectors from another, we must consider

[1] [*Non zero.*]

[2] [*By the second assumption.*]

[3] [Or collinear.]

[4] ["The tensor of the quotient."]

not only its *magnitude* (or *quantity*), but also its PLANE: since otherwise, in violation of the principle[1]..., we should have $OB'{:}OA = OB{:}OA$, if OB and OB' were *two distinct rays* or sides of a *cone* of revolution, with OA for its *axis;* in which case...they would necessarily be *unequal vectors.* For a similar reason, we must attend also to the *contrast* between two *opposite angles*, of equal magnitudes, and in one *common plane.* In short, for the purpose of knowing *fully* the *relative direction* of two co-initial lines OA, OB in *space*, we ought to know not only *how many degrees*...the *angle* AOB contains; but also...the *direction of the rotation* from OA to OB: including a knowledge of the *plane, in which* the rotation is performed; and of the *hand* (as *right* or *left*, when *viewed* from a known *side* of the plane), *towards which* the rotation is *directed.*

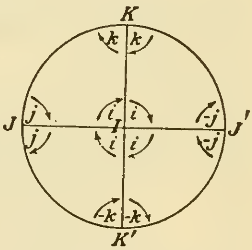
"111. Or, if we agree to *select* some one *fixed hand* (suppose the *right* hand), and to call all *rotations positive* when they are directed towards *this* selected hand, but all rotations *negative* when they are directed towards the *other hand*, then, for *any given angle* AOB, supposed for simplicity to be less than two right angles, and considered as representing a *rotation in a given plane* from OA to OB, we may speak of *one perpendicular* OC *to that plane* AOB as being the *positive axis* of that rotation; and of the *opposite perpendicular* OC′ to the same plane as being the *negative axis* thereof: the rotation around the positive axis being *itself* positive, and *vice-versa.* And then the *rotation* AOB may be considered to be entirely *known*, if we know, Ist, its *quantity*, or the *ratio* which it bears to a *right rotation;* and IInd, the *direction* of its *positive axis*, OC, but not without knowledge of these *two* things, or of some data equivalent to them. But whether we consider the *direction of an* AXIS, or the *aspect of a* PLANE, we find (as indeed is well known) that the *determination* of such a *direction*, or of such an *aspect*, depend on TWO *polar coordinates*, or other *angular elements.*

"112. It appears, then, from the foregoing discussion, that *for the complete determination*, of what we have called the *geometrical* QUOTIENT *of two coinitial Vectors*, a *System of Four Elements*, admitting each separately of numerical expression, *is generally required.* Of these four elements, *one* serves to determine the *relative length* of the two lines compared; and the other *three* are in general necessary, in order to determine *fully* their *relative direction.* Again, of these three latter elements, *one* represents the

[1] [Assumption 2.]

mutual *inclination*, or *elongation*, of the two lines; or the *magnitude* (or quantity) of the *angle* between them; while the *two others* serve to determine the *direction* of the *axis*, perpendicular to their common *plane*, round which a *rotation* through that angle is to be performed, in a *sense* previously selected as the *positive* one (or towards a fixed and previously selected *hand*), for the purpose of *passing* (in the simplest way, and therefore in the plane of the two lines) from the *direction* of the *divisor-line*, *to* the direction of the *dividend-line*. And *no more than four* numerical *elements* are necessary for our present purpose: because the *relative length* of two lines is not changed when their lengths are altered proportionally, nor is their *relative direction* changed, when the *angle* which they form is merely *turned* about, *in its own plane*. On account, then, of this *essential connexion* of that *complex relation* between two lines, which is *compounded* of a *relation of lengths*, and of a *relation of directions*, and to which we have given (by an *extension* from the theory of *scalars*), the name of a *geometrical quotient*, with a *System of* FOUR *numerical Elements*, we have already a *motive* for saying that '*The Quotient of two Vectors is generally a Quaternion*'."[1]

"181. Suppose that OI, OJ, OK are any three given and coinitial but rectangular unit lines, the rotation around the first from the second to the third being positive; and let OI′, OJ′, OK′ be the three unit vectors respectively opposite to these, so that



$$OI' = -OI, OJ' = -OJ, OK' = -OK.$$

Let the three new symbols i, j, k denote a *system of three right versors*,[2] *in three mutually rectangular planes*, . . . ; so that . . . i = OK:OJ, j = OI:OK, k = OJ:OI, as the figure may serve to illustrate. We shall then have these other expressions for the same three versors.

$$\begin{aligned} i &= OJ':OK = OK':OJ' = OJ:OK'; \\ j &= OK':OI = OI':OK' = OK:OI'; \\ k &= OI':OJ = OJ':OI' = OI:OJ'; \end{aligned}$$

---

[1] ["Quaternion . . . signifies . . . a Set of Four."]

[2] [A right versor is an operator which produces a rotation of a right angle about a given axis in a given direction.]

while the three respectively *opposite* versors may be thus expressed:

$$-\mathrm{i} = OJ:OK = OK':OJ = OJ':OK' = OK:OJ'$$
$$-\mathrm{j} = OK:OI = OI':OK = OK':OI' = OI:OK'$$
$$-\mathrm{k} = OI:OJ = OJ':OI = OI':OJ' = OJ:OI'$$

and from the comparison of these different expressions several important symbolical consequences follow...

"182. In the *first* place, since

$$\mathrm{i}^2 = (OJ':OK).(OK:OJ) = OJ':OJ, \text{ etc.},$$

we deduce the following equal values for the *squares* of the new symbols:

I $$\mathrm{i}^2 = -1;\ \mathrm{j}^2 = -1;\ \mathrm{k}^2 = -1,$$

. . . .

In the *second* place, since

$$\mathrm{ij} = (OJ:OK').(OK':OI) = OJ:OI, \text{ etc.},$$

we have the following values for the *products* of the same three symbols, or versors, when taken *two by two*, and in a certain *order* of *succession*...:

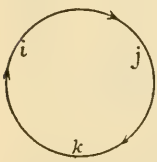
II $$\mathrm{ij} = \mathrm{k};\ \mathrm{jk} = \mathrm{i};\ \mathrm{ki} = \mathrm{j}.$$

But in the *third* place..., since

$$\mathrm{ji} = (OI:OK).(OK:OJ) = OI:OJ, \text{ etc.},$$

we have these other and *contrasted* formulae, for the *binary products* of the *same* three right versors, when taken as factors with an *opposite order:*

III $$\mathrm{ji} = -\mathrm{k};\ \mathrm{kj} = -\mathrm{i};\ \mathrm{ik} = -\mathrm{j}.$$



Hence, while the *square of each* of the *three right versors*, denoted by these *three new symbols*, i, j, k, is equal to *negative unity*, the *product of any two* of them is equal either to the *third itself*, or to the *opposite* of that third versor, according as the *multiplier precedes* or *follows* the *multiplicand*, in the *cyclical succession*

$$\mathrm{i}, \mathrm{j}, \mathrm{k}, \mathrm{i}, \mathrm{j}, \ldots$$

which the annexed figure may give some help towards remembering.

"183. Since we have thus $\mathrm{ji} = -\mathrm{ij}$, ...we see that the *laws of combination of the new symbols*, i, j, k, are *not in all respects the same* as the corresponding laws in *algebra;* since the *Commutative Property of Multiplication*, or the *convertibility* of the places of the *factors*

without change of value of the *product*, does *not here* hold good; which arises from the circumstance that the factors to be combined are here diplanar versors. It is therefore important to observe that there *is* a respect in which the *laws of* i, j, k *agree* with usual and *algebraic laws:* namely, in the Associative Property of Multiplication; or in the property that the new symbols always obey the associative formula

$$\iota\kappa\lambda = \iota\kappa\lambda,$$

whichever of them may be substituted for ι, for κ, and for λ; in virtue of which equality of values we may *omit the point* in any such symbol of a *ternary product* (whether of equal or unequal factors), and write it simply as $\iota\kappa\lambda$. In particular, we have thus,

$$\mathrm{i.jk} = \mathrm{i.i} = \mathrm{i}^2 = -1; \qquad \mathrm{ij.k} = \mathrm{k.k} = \mathrm{k}^2 = -1;$$

or briefly

$$\mathrm{ijk} = -1.$$

We may, therefore,. . .establish the following important *Formula:*

$$\mathrm{i}^2 = \mathrm{j}^2 = \mathrm{k}^2 = \mathrm{ijk} = -1;$$

. . .which we shall find to contain (virtually) *all the laws of the symbols* i, j, k, and therefore to be a *sufficient symbolical basis* for the whole *Calculus of Quaternions:* because it will be shown that *every quaternion can be reduced to the Quadrinomial Form,*

$$\mathrm{q} = \mathrm{w} + \mathrm{ix} + \mathrm{jy} + \mathrm{kz},$$

where w, x, y, z compose a *system of four scalars*, while i, j, k are the same *three right versors as above*."

"*If two right versors in two mutually rectangular planes, be multiplied together in two opposite orders, the two resulting products will be two opposite right versors, in a third plane, rectangular to the two former;* or in symbols. . .

$$\mathrm{q'q} = -\mathrm{qq'}$$

. . .In *this* case, therefore, we have what would be in algebra a *paradox*. . .When we come to examine what, in the last analysis, may be said to be the *meaning* of this last equation, we find it to be simply this: that *any two quadrantal* or *right rotations, in planes perpendicular to each other, compound themselves into a third right rotation, as their resultant in a plane perpendicular to each of them:* and that this *third* or *resultant* rotation has one or other of two *opposite directions*, according to the *order* in which the two *component rotations* are taken, so that one shall be *successive* to the other."

GRASSMANN

On the Ausdehnungslehre

(Translated from the German by Dr. Mark Kormes, New York City.)

Hermann Günther Grassmann (1809–1877) was professor at the gymnasium in Stettin. In 1844 he published *Die lineale Ausdehnungslehre, ein neuer Zweig der Mathematik...*, which was not generally understood on account of its abstract philosophical form. Grassmann therefore rewrote his book many years later and published it in Berlin (1862) under the title *Die Ausdehnungslehre, Vollständig und in strenger Form bearbeitet.* Besides many other important contributions to mathematics he distinguished himself as a scholar of Sanskrit literature.

In his *Ausdehnungslehre* Grassmann created a symbolic calculus so general that its definitions and theorems can be easily applied not only to geometry of n dimensions but also to almost every branch of mathematics. This calculus forms the basis of vector analysis. By its aid Grassmann derived fundamental theorems on determinants and solved many elimination problems in a most elegant manner. In connection with the problem of Pfaff and the theory of partial differential equations, his theorems are of great importance.

The following translation is limited to (1) the development of the idea of non-commutative multiplication—the combinatory (outer) and the inner products; and (2) certain passages relating to geometry. It is taken from the 1862 edition,—the one which made Grassmann's influence felt.

1. We say that a quantity a is derived from the quantities b, c...by means of numbers β, γ,...if

$$a = \beta b + \gamma c + \ldots$$

where β, γ...are real numbers, rational or irrational, and may be equal to zero. We also say in such a case that a is derived numerically from b, c...

2. We further say that two or more quantities a, b, c...are numerically related if one of them can be derived numerically from the others; for example:

$$a = \beta b + \gamma c + \ldots$$

where β, γ,...are real numbers...

3. A quantity from which a set of other quantities may be derived numerically is called a *unit* and in particular a unit which cannot be derived numerically from any other unit is called a *primitive unit.* The unit of numbers is called the *absolute unit,*

all other units are *relative*. Zero should never be regarded as a unit.

5. An *extensive* quantity is any expression derived by means of numbers from a system of units[1] (which system should not consist solely of the absolute unit). The numbers used are called *derivation-coefficients*. For example the polynominal

$$\alpha_1 e_1 + \alpha_2 e_2 + \ldots \text{ or } \Sigma \alpha e \text{ or } \Sigma \alpha_r e_r$$

is an extensive quantity when $\alpha_1, \alpha_2, \ldots$ are real numbers and $e_1, e_2 \ldots$ is a system of units. Only if the system of units consists solely of the absolute unit (1) the derived quantity is not an extensive but a numerical quantity...

6. To add two extensive quantities derived from the same system of units we add the derivation-coefficients of the corresponding units:

$$\Sigma \alpha_r e_r + \Sigma \beta_r e_r = \Sigma(\alpha_r + \beta_r) e_r$$

. . . .[2]

9. All laws of algebraic addition and subtraction hold for the extensive quantities...

10. To multiply an extensive quantity by a number we multiply all its derivation-coefficients by this number:

$$\Sigma \alpha_i e_i . \beta = \beta . \Sigma \alpha_i e_i = \Sigma(\alpha_i \beta) e_i$$

. . . .[3]

13. All laws of algebraic multiplication and division hold for the multiplication and division of an extensive quantity by a number.

37. A product $[ab]$ of two extensive quantities a and b is defined as an extensive quantity (or a numerical quantity) obtained in the following way: Multiply each of the units from which the first quantity a is derived by each of the units from which the second quantity b is derived so that the unit of the first quantity is always the first factor and the unit of the second quantity is the second factor of the product; multiply then every such product by the product of the corresponding derivation-coefficients and add all the products so obtained:

$$[\Sigma \alpha_r e_r \Sigma \beta_s e_s] = \Sigma \alpha_r \beta_s [e_r e_s] \ldots$$

---

[1] [A set of units not related numerically (See 4 in the original text).]

[2] [Here follows a similar definition for subtraction. (See 7 in original text.)]

[3] [Here follows a similar definition for division by a number. (See 11 in the original text.)]

Remark. Inasmuch as according to the above definition the product of extensive quantities is either an extensive or a numerical quantity, we must be able (see 5) to derive it numerically from a system of units. What is this system of units and how are we to derive numerically the products $e_r e_s$ from it, is not explained in the definition. Therefore if we are to determine exactly the concept of a particular product we must agree upon certain necessary rules... Consider, *e.g.*, the product $P = [(x_1e_1 + x_2e_2)(y_1e_1 + y_2e_2)] = x_1y_1[e_1e_1] + x_1y_2[e_1e_2] + x_2y_1[e_2e_1] + x_2y_2[e_2e_2]$... We could then agree that the four products $[e_1e_1]$, $[e_1e_2]$, $[e_2e_1]$ and $[e_2e_2]$ constitute the system of units from which P is to be derived numerically so that the numbers x_1y_1, x_1y_2, x_2y_1, and x_2y_2 are the derivation-coefficients. We would have then a particular product characterized by the fact that no equations are necessary for its determination. We could on the other hand select three of them; $[e_1e_1]$, $[e_1e_2]$ and $[e_2e_2]$ as units and agree that $[e_2e_1] = [e_1e_2]$, the derivation-coefficients of P would be then: x_1y_1, $(x_1y_2 + x_2y_1)$ and x_2y_2; this kind of a product is characterized by the fact that the laws which govern it are identical with those of algebraic multiplication. We could also select $[e_1e_2]$ as a unit and agree that $[e_1e_1] = 0$, $[e_2e_1] = -[e_1e_2]$, and $[e_2e_2] = 0$; in this case the product P would have only one derivation-coefficient, namely $x_1y_2 - x_2y_1$. Such products are subsequently designated as *combinatory*. We may finally agree to select a system of units not containing any one of the products $[e_1e_1]$, $[e_1e_2]$, $[e_2e_1]$, $[e_2e_2]$, and then to determine how to derive these four products from this system; *e.g.*, we may choose as the system of units the absolute unit and agree that $[e_1e_1] = 1$, $[e_1e_2] = 0$, $[e_2e_1] = 0$, $[e_2e_2] = 1$. Under such conditions P becomes a numerical quantity namely $P = x_1y_1 + x_2y_2$. These products are subsequently designated as *inner products*.

50. Every multiplication for which the determining equations[1] remain true if we substitute in place of the units quantities numerically derived from them is said to be a *linear* multiplication.

51. Besides the multiplication which has no determining equation or the multiplication for which all products are zero there exist only two kinds of linear multiplication: the determining equation for the one is

$$[e_r e_s] + [e_s e_r] = 0 \qquad (1)$$

[1] [A numerical relation between the products of units. (See 48 in the original text.)]

and for the other

(2) $$[e_r e_s] = [e_s e_r] \ldots$$

52. A *combinatory product* is defined as a product the factors of which are derived from a system of units, provided that the sum of any two products of units obtained from each other by interchanging the last two factors is equal to zero, while every product consisting solely of different units is not zero. The factors of this product are called *simple factors*. If b and c are units and A is an arbitrary set of units the above condition is expressed in the following form:

$$[Abc] + [Acb] = 0.$$

53. In every combinatory product we may interchange the two last factors provided that we change the sign of the product, or $[Abc] + [Acb] = 0$, in the case when A is an arbitrary set of factors and b and c are simple factors.

Proof.—1. Suppose that b and c are units. Since A is a set of arbitrary factors and since these factors may be derived numerically from units we obtain after substitution an expression for A of the form: $A = \Sigma\alpha_r E_r$, where E_r are products of units. Thus we have

$$\begin{aligned}[Abc] + [Acb] &= [\Sigma\overline{\alpha_r E_r}bc] + [\Sigma\overline{\alpha_r E_r}cb] \\ &= \Sigma\alpha_r[E_r bc] + \Sigma\alpha_r[E_r cb)] \\ &= \Sigma\alpha_r([E_r bc] + [E_r cb]) \\ &= \Sigma\alpha_r 0 \qquad (52) \\ &= 0.\end{aligned}$$

2. Suppose now that b and c are numerically derived from the units $e_1, e_2, \ldots$, *e. g.*, $b = \Sigma\beta_r e_r$, $c = \Sigma\gamma_r e_r$ we have then:

$$\begin{aligned}[Abc] + [Acb] &= [A\Sigma\beta_r e_r \Sigma\gamma_r e_r] + [A\Sigma\gamma_r e_r \Sigma\beta_r e_r] \\ &= \Sigma\beta_r\gamma_s[Ae_r e_s] + \Sigma\gamma_s\beta_r[Ae_s e_r] \\ &= \Sigma\beta_r\gamma_s([Ae_r e_s] + [Ae_s e_r]) \\ &= \Sigma\beta_r\gamma_s 0 \qquad \text{(proof 1.)} \\ &= 0.\end{aligned}$$

55. Any two factors of a combinatory product may be interchanged provided that we change the sign of the product:

$$P_{a,b} = -P_{b,a} \text{ or } P_{a,b} + P_{b,a} = 0.$$

. . . .

60. If two simple factors of a combinatory product are equal the product is equal to zero,

$$P_{a,a} = 0.$$

Proof...According to 55 we have: $P_{a,b} + P_{b,a} = 0$ hence if we put $a = b$,

$$P_{a,a} + P_{a,a} = 0 \text{ or } 2P_{a,a} = 0;$$

thus

$$P_{a,a} = 0.$$

61. A combinatory product is equal to zero if its simple factors are numerically related,—*i. e.*, $[a_1a_2a_3\dots a_m] = 0$,—if one of the quantities $a_1a_2\dots a_m$ can be derived numerically from the others; *e.g.*, if

$$a_1 = \alpha_2a_2 + \alpha_3a_3 + \dots + \alpha_ma_m.$$

Proof: If we substitute in the product for a_1 the above expression we obtain

$$\begin{aligned}[a_1a_2a_3\dots a_m] &= [(\alpha_2a_2 + \alpha_3a_3 + \dots + \alpha_ma_m)a_2a_3\dots a_m] \\ &= \alpha_2[a_2a_2a_3\dots a_m] + \alpha_3[a_3a_2a_3\dots a_m] + \dots \\ &\qquad + \alpha_m[a_ma_2a_3\dots a_m] \\ &= \alpha_20 + \alpha_3.0 + \dots + \alpha_m.0 \\ &= 0.\end{aligned}$$

63. In order to obtain a combinatory product of n simple factors which are derived numerically from the n quantities $a_1, a_2, \dots a_n$ we form the determinant of the derivation-coefficients where the coefficients of the first factor form the first row, etc., and we multiply this determinant by the combinatory product of the quantities $a_1, a_2\dots a_n$:

$$\begin{aligned}[(\alpha_1{}^{(1)}a_1 + \dots + \alpha_n{}^{(1)}a_n)(\alpha_1{}^{(2)}a_1 + \dots + \alpha_n{}^{(2)}a_n)\dots(\alpha_1{}^{(n)}a_1 + \dots \\ + \alpha_n{}^{(n)}a_n)] = \Sigma \mp \alpha_1{}^{(1)}\alpha_2{}^{(2)}\dots\alpha_n{}^{(n)}[a_1a_2\dots a_n]^1\dots\end{aligned}$$

68. All laws of combinatory multiplication hold if we substitute for the n primitive units an arbitrary set of n quantities derived numerically from these units provided that these n quantities are not related numerically.

[1] [The symbol $\Sigma \mp \alpha_1{}^{(1)}\alpha_2{}^{(2)}\dots\alpha_n{}^{(n)}$ is used by Grassmann for the determinant

$$\begin{vmatrix} \alpha_1{}^{(1)}\alpha_2{}^{(1)}\dots\alpha_n{}^{(1)} \\ \alpha_1{}^{(2)}\alpha_2{}^{(2)}\dots\alpha_n{}^{(2)} \\ \dots\dots\dots\dots\dots \\ \alpha_1{}^{(n)}\alpha_2{}^{(n)}\dots\alpha_n{}^{(n)} \end{vmatrix}$$

64. Multiplicative combinations from a set of quantities are defined as combinations without repetition of these quantities whereby each combination is a combinatory product, the factors of which are the elements of the combination; *e. g.*, the multiplicative combinations of second class from the three quantities a, b, c are: $[ab]$, $[ac]$ and $[bc]$.

77. The multiplicative combinations of class m of the primitive units shall be called *units of order m*. A quantity derived numerically from units of order m shall be called a *quantity of order m*. Such quantity is said to be *simple* if it may be represented as a combinatory product of quantities of the first order, otherwise it is called a *composite* quantity. The aggregate of all quantities which can be derived numerically from the simple factors of a simple quantity is called the *domain* of this quantity.

77*b*...Remark. As an example of a composite quantity let us consider the sum $(ab) + (cd)$ where a, b, c, d are quantities not related numerically. Suppose $(ab) + (cd)$ were a simple quantity, *e.g.*, equal to $(p.q)$; we would have $[(ab + cd)(ab + cd)] = [pqpq] = 0$. But $[(ab + cd)(ab + cd)] = [abcd] + [cdab]$ on account of $[abab]$ and $[cdcd]$ being equal to zero. It is however $[abcd] = [cdab]$. Hence $[(ab + cd).(ab + cd)] = 2.[abcd]$. If $(ab) + (cd)$ were a simple quantity $[abcd]$ would be equal to zero and a, b, c, d would be related numerically which contradicts the assumption.

78. The *outer* product of two units of higher order is defined as the combinatory product of the simple factors of those quantities, whereby the arrangement of these factors remains undisturbed:

$$[(e_1e_2\ldots e_m)(e_{m+1}\ldots e_n)] = [e_1e_2\ldots e_n]$$

NOTE.—The name *outer multiplication* is used to emphasize the fact that this product holds if and only if one factor is entirely outside of the domain of the other factor.

79. In order to obtain the outer product of two quantities A and B we form the combinatory product of the simple factors of the first quantity with those of the second quantity:

$$[(ab\ldots).(cd\ldots)] = [ab\ldots cd\ldots].$$

80. The parentheses have no effect on the outer product:

$$[A(BC)] = [ABC]\ldots$$

83. Given a sum of simple quantities S and a set of m quantities of the first order $a_1a_2\ldots a_m$ which are not related numerically.

If the outer product of S with every one of the quantities $a_1, a_2 \ldots a_m$ is equal to zero, S may be represented as an outer product in which $a_1, a_2 \ldots a_m$ are factors; *i. e.*,

$$S = [a_1, a_2 \ldots a_m S_m)$$

if

$$0 = [a_1 S] = [a_2 S] = \ldots = [a_m S] \ldots$$

86. The *principal* domain is the domain of the primitive units from which the quantities under consideration were derived...

89. Let us consider a principal domain of order n and let us assume that the product of the primitive units is equal to 1. If E is a unit of any arbitrary order (*i. e.*, either one of the primitive units or a product of a number of them) the *complement* of E is defined as a quantity which is equal to the combinatory product E' of all units which are not in E. The complement is positive or negative according to whether $[EE']$ is equal to $+1$ or to -1. We will denote the complement by a vertical line before the given quantity. The complement of a number shall be this number itself. Thus we have: $|E = [EE']\, E'$ if E and E' contain all the units $e_1 e_2 \ldots e_n$, and if $[e_1 e_2 \ldots e_n] = 1$; also $|\alpha = \alpha$ when α is a number.

90. The complement of an arbitrary quantity A is the quantity $|A$ obtained if in the expression for A we substitute the complements in the place of the units; *i. e.*,

$$|(\alpha_1 E_1 + \alpha_2 E_2 + \ldots) = \alpha_1 |E_1 + \alpha_2 |E_2 + \ldots$$

where $E_1 E_2 \ldots$ are units of any order whatsoever.

91. The outer product of a unit and its complement is equal to 1:

$$[E|E] = 1.$$

Proof.—If E' is the combinatory product of all primitive units not contained in E we have (according to 89):

$$|E = \mp E' \text{ according to whether } [EE'] = \mp 1.$$

In the case of the lower sign we have: $[E|E] = [EE'] = 1$ and in the case of the upper sign: $[E|E] = -[EE'] = -(-1) = 1$.

92. The complement of a complement of a quantity A is either equal to A or to $-A$ according to whether the product of the orders of A and of its complement is even or odd; *i. e.*,

$$||A = (-1)^{qr} A$$

if q and r are the orders of A and $|A$ respectively...

Note.—If both r and q are odd, as for example in the case of a domain of order two and complements of quantities of order one,

we have $||A = -A$ so that in such a case the symbol $|$ obeys the same laws as $i = \sqrt{-1}$, which gives an interpretation of the imaginary in the real domain...

93. If the principal domain is of order n and if n is odd we have

$$||A = A;$$

if n is even

$$||A = (-1)^q A$$

where q is the order of A...

94. If the sum of the orders of two units is equal to or less than n,—*i. e.*, the order of the principal domain,—the outer product of these units is called a *progressive* product provided that the progressive product of the primitive units is equal to 1. If on the other hand the sum of the orders of two units is greater than n, the *regressive* product of these units is given by a quantity whose complement is equal to the progressive product of the complements of these units. We shall refer to both the regressive and progressive products as *relative products*...

97. The product of the complements of two quantities is equal to the complement of the product of these quantities:

$$[|A|B] = |[AB]\ldots$$

...If the product of two quantities is progressive, the product of their complements is regressive provided that we agree to consider the product of order zero as a progressive and as a regressive product at the same time...

122. A *mixed product*[1] of three quantities $[ABC]$ is equal to zero if and only if either $[AB] = 0$, or all the quantities A, B, and C are contained in a domain of order less than n, or the quantities A, B, and C have a domain of order more than 0 in common...

137. The *inner product* of two units of an arbitrary order is defined as the relative product of the first unit and the complement of the second unit; *i. e.*, if E and F are units of an arbitrary order the inner product is given by $[E|F]$.

138. The inner product of two quantities A and B is equal to the relative product of the first quantity and the complement of the second quantity, *i. e.*, $[A|B]$...

139. If the factors of an inner product are of orders α and β and if the principal domain is of order n, the inner product is of order

[1] [That is, a product in which both the progressive and the regressive multiplication is used.]

$n + \alpha - \beta$ or $\alpha - \beta$ according to whether β is greater than α or not...

141. The inner product of two quantities of equal order is a number.

Proof. The difference between the orders is then zero, hence the inner product is of order zero,—*i. e.*, a number.

142. The inner product of two equal units is 1, the inner product of two different units of equal order is zero; *i. e.*,

$$[E_r|E_r] = 1, \; [E_r|E_s] = 0\ldots$$

143. If $E_1E_2\ldots E_m$ are units of an arbitrary but equal order we have:

$$[(\alpha_1E_1 + \alpha_2E_2 + \ldots + \alpha_mE_m)|(\beta_1E_1 + \ldots + \beta_mE_m)] = \alpha_1\beta_1 + \alpha_2\beta_2 + \ldots + \alpha_m\beta_m\ldots$$

144. The two factors of an inner product may be interchanged provided they are of the same order; *i. e.*,

$$[A|B] = [B|A].$$

Proof.—If $E_1\ldots E_m$ are the units and $A = \Sigma\alpha_rE_r$, $B = \Sigma\beta_sE_s$, we have from 143,

$$[A|B] = \Sigma\alpha_r\beta_r = \Sigma\beta_r\alpha_r = [B|A].$$

145. For the sake of simplicity we write

$$[A|A] = A^2,$$

and we call it the *inner square* of A...

147. The inner product of two units E and F is not equal to zero if and only if one of the units is incident[1] with the other...

148. If E and F are units and $[EF] \neq 0$, we have

$$[EF|E] = F \text{ and } [F|EF] = |E\ldots$$

151. The *numerical value* of a quantity A is defined as the *positive square root* of the inner square of that quantity. Two quantities are said to be numerically equal if their numerical values, *i. e.*, their inner squares,—are equal.

152. Two quantities different from zero are said to be *orthogonal* if their inner product is zero...

153. A set of n numerically equal quantities of the first order which are orthogonal to each other is called an *orthogonal system*

[1] [A quantity is said to be *incident* with another if its domain is incident; that is all quantities of the domain of the first quantity are also quantities of the domain of the second quantity, but not vice versa.]

of order n; in the case that the domain is also of order n we speak of a *complete orthogonal system.* The numerical value of the given quantities is said to be the numerical value of the orthogonal system. Every orthogonal system having the numerical value 1 is said to be *simple*...

157. The quantities of an orthogonal system are not related numerically and every quantity of the first order can be derived numerically from any arbitrary complete orthogonal system...

162. The system of the primitive units is a complete orthogonal system, whose numerical value is 1.

Proof.—If $e_1 e_2 \ldots e_n$ are the primitive units, we have

$$1 = e_1^2 = e_2^2 = \ldots = e_n^2$$
$$0 = [e_1|e_2] = \ldots$$

163. In every domain of order m we can establish an orthogonal system of order m having an arbitrary numerical value so that this system is a part of the complete orthogonal system...

168. All previous theorems[1] remain true if we replace the system of the primitive units by an arbitrary complete orthogonal system which has the numerical value 1...

175. Given two quantities A and B of order m each of which is composed of m simple factors. The inner product of these two quantities is equal to the determinant of m rows and m columns which is obtained by forming inner products of every simple factor of one quantity with every simple factor of the other quantity; *i.e.*

$$[abc\ldots|a'b'c'\ldots] = \text{Determ.} \begin{vmatrix} [a|a'], & [a|b'], & [a|c'], \ldots \\ [b|a'], & [b|b'], & [b|c'], \ldots \\ [c|a'], & [c|b'], & [c|c'], \ldots \\ \ldots & \ldots & \ldots \end{vmatrix}$$

. . . .

216. Given a point E and let us assume that three lines of equal length and perpendicular to each other are the principal units. If $\alpha_1, \alpha_2, \alpha_3$ are arbitrary numbers the expression:

$$(a) \qquad E + \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3$$

defines the point A, obtained in the following manner: From E we proceed along the segment EB which is equal to $\alpha_1 e_1$, that is, which has the same direction as e_1 or the opposite direction according to whether α_1 is positive or negative and the distance EB is

[1] [Relative to the inner product.]

in the same ratio to e_1 as α_1 is to 1. We then proceed from B along the segment BC equal to $\alpha_2 e_2$ in the above sense and finally from C we proceed along the segment CA which is equal to $\alpha_3 e_3$ in the same manner. Furthermore the expression:

$$(b) \qquad \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3$$

defines a segment, that is, a straight line of a given length and directions and namely such a particular segment which has the same length and direction as the line connecting the point E with the point

$$E + \alpha_1 e_1 + \alpha_2 e_2 + \alpha_3 e_3 \ldots$$

229. Every segment of the space may be derived numerically from three arbitrary segments which are not parallel to a plane...

231. If three segments are related numerically they are parallel to a plane...

232. All points of the space can be derived numerically from four arbitrary points which do not lie in one plane...

234. Every point of a straight line may be derived numerically from two arbitrary points of this line...

235. If three points are related numerically they lie in a straight line...

236. If four points are related numerically they lie in a plane...

237. In the space a domain of first order is a point, the domain of second order the unlimited straight line, that of third order the unlimited plane and that of fourth order the unlimited space.

245. The combinatory product of two points vanishes if and only if the two points coincide; the combinatory product of three points vanishes if the points lie in a straight line, that of four points if they lie in a plane and the combinatory product of five points always vanishes...

249. The product $[AB]$ shall be called a *segment* and we shall say that it is a part of the unlimited line AB and that it is of equal length and direction as the segment AB...

273. The sum of two finite segments the lines of which intersect is a segment and its line passes through the point of intersection of the other two lines; the direction and length of this segment are the same as those of the diagonal of the parallelogram formed by segments of the same length and direction as the two original segments...

288. Planimetric multiplication is defined as the relative multiplication with respect to the plane; stereometric multiplication

as relative multiplication with respect to the space (as domain of order four)...

306.[1] The equation of a point x which lies in a straight line with the points a and b is given by:

$$[xab] = 0.$$

Proof.—$[xab]$ vanishes (according to 245) if and only if x lies in a straight line with a and b...

307. The equation of a straight line X which passes through the same point as the straight lines A and B is given by

$$[XAB] = 0\ldots$$

309. If $P_{n,x}$ is a planimetric product of order zero in which the point x is contained n times and if the other factors are only fixed points or lines, the equation:

$$P_{n,x} = 0$$

is then the point-equation of an algebraic curve of order n, provided that it is not satisfied by every point x...

310. If $P(n, X)$ is a planimetric product of order zero in which the line X is contained n times and as the other factors are only fixed points or lines, the equation

$$P(n, X) = 0$$

is then the line-equation of an algebraic curve of class n...

311. If $P_{n,x}$ is a stereometric product of order zero, which contains the point x n times and as other factors has only fixed points, lines or planes, the equation:

$$P_{n,x} = 0$$

is the point-equation of an algebraic surface of order n...

323. The equation of a conic section, passing through the five points a, b, c, d, e, no three of which lie in a straight line is given by...

$$[xaBc_1.Dex] = 0$$

where $B = [cd]$ $c_1 = [ab.de]$ $D = [b.c]$...

324. If A, B, C are three straight lines in space no two of which intersect then:

$$[xABCx] = 0$$

[1] [From now on Grassmann uses small letters to denote points and capital letters to denote lines.]

is the equation of the surface of second order which contains the three straight lines A, B, C...

330. For the purpose of inner multiplication we shall always assume as principal units three segments of equal length and perpendicular to each other ($e_1e_2e_3$), in the plane two such segments (e_1e_2) and we shall assume that the length of these segments shall be the unit of length, $[e_1e_1e_2]$ the unit of volume and $[e_1e_2]$ the unit of area.

331. For the plane[1] the concept of length coincides with the concept of numerical value, orthogonal means perpendicular...

[1] [Also for the space (see 333 of original text).]

Index